"Elements of the Mathematic"
Two interesting examples at the intersection of math

By
Hunter Davis

Senior Honors Thesis
Deparment of Mathematics
University of North Carolina at Chapel Hill

March 24, 2021

Approved: ______________________________
Idris Assani, Thesis Advisor
Hans Christianson, Reader
Mark Williams, Reader

Acknowledgements

I would like to thank the entire UNC Math department for providing an exciting and rigorous environment to pursue mathematics. At every step of my journey here I have felt welcome, encouraged, and challenged. Specifically, I would also like to thank my readers, Professors Christianson and Williams.

Finally, I owe a tremendous amount of gratitude to my thesis advisor Professor Assani, both for providing invaluable feedback and assistance throughout this project as well as over the course of my entire UNC career.

Contents

1 Introduction

In 1939, the first books of the Nicolas Bourbaki group's *Éléments de Mathématique* were published. While those works contain some very interesting mathematics, the most immediately interesting aspect of them is the singular noun in the title- Mathematique instead of *Mathematiques.* In what ways can the study of math be considered a singular discipline as opposed to a collection of disparate topics?

This honors project attempts to provide an undergraduate-level exposition of two important results which connect seemingly independent fields of mathematics. First, we demonstrate how a very un-intuitive result from elementary set theory can play a crucial role in proving a result about the description of the Borel Sets of R^n. In the second part, we walk through and expand upon proofs from Guy Robin (16) and Jeffrey Lagarias (17) which demonstrate that the Riemann Hypothesis is equivalent to a seemingly elementary inequality.

2 Part 1: Set Theory and the Groundwork of Mathematics

In their introduction to the first book in their series, Nicolas Bourbaki writes "in the past it was it was thought that every branch of mathematics depended on its own particular intuitions which provided its concepts and prime truths, nowadays it is known to be possible, logically speaking, to derive practically the whole of mathematics from a single source, the Theory of Sets." (1) Since their exposition, several other attempts have been made to demonstrate how Set Theory underlies most of mathematics, and to show just how far results from elementary Set Theory can go in terms of proving very sophisticated results. In this paper, we will attempt to summarize one particularly good attempt by Paul Halmos in his *Naive Set Theory* (2), as well take it further by applying his one of culminating results, transfinite induction, to demonstrate the fundamental result that the Borel Hierarchy exhausts the smallest σ-algebra containing the open sets.

3 Summary of Halmos's *Naive Set Theory*

3.1 Chapter 1: Axiom of Extension

Here Halmos introduces some of the basic concepts necessary to understand set theory, such as the idea of 'belonging', denoted by $\in$, as well as equality, denoted by $=$. However, while the relationship of 'belonging' is simply accepted as axiomatic, equality is specifically defined as sets having the same elements (this is the Axiom of Extension).

He also defines the $\subset$ relationship, defined as $A \subset B \iff a \in A \to a \in B$.

Theorem 3.1 *The $\subset$ relationship is transitive, antisymmetric, and reflexive.*

1. We aim to show that $A \subset B \wedge B \subset C \to A \subset C$. By our definitions, $a \in A \implies a \in B$. But $a \in B \to a \in C$, therefore $a \in A \to a \in C \Rightarrow A \subset C$. And so the relationship is transitive.

2. We aim to show that if $A \neq B$ and $A \subset B$, then $B \not\subset A$. Observe that if $a \in A$ then $a \in B$ by hypothesis. So if $B \subset A$ then $b \in B$ would imply $b \in A$, and so $a \in A \iff a \in B$ and therefore by the Axiom of Extension, $A = B$ which is a contradiction with our initial assumption. Therefore $\subset$ is anti-symmetric.

3. Finally, we aim to show $A \subset A$, which is trivially true since $a \in A \to a \in A$. So $\subset$ is reflexive.

3.2 Chapter 2: Axiom of Specification

In this chapter Halmos begins by defining the logical syntax he is using for the book. The valid symbols he takes as given (although he spells them out instead of giving the standard symbols), are $\in, =, \neg, \vee, \wedge, \implies, \iff, \exists, \forall$. Having defined these, he is prepared to define the Axiom of Specification, which states that any sentence constructed from these operators that has a free variable can be used to define a subset out of a larger set.

Theorem 3.2 *There does not exist a set which contains everything. Ie.* $\nexists A : \{x : x = x\}$

Take an arbitrary set A and use the axiom of specification letting $B_A = \{x \in A : x \notin x\}$. Then if $B \in A$, either $B_A \in B_A \implies B_A \notin B_A$, or $B_A \notin B_A \implies B_A \in B_A$. Since both of these are contradictions, it must mean $B \notin A$. Therefore for any set A, there exists at least one element, B_A as defined above, which is not contained in it.

Halmos then talks about why this result justifies the use of the Axiom of Specification (and specifically justifies it demanding a set A from which to separate), since it avoid Russel's Paradox.

3.3 Chapter 3: Unordered Pairs

Here Halmos starts off by assuming that there exists a set, from which it's fairly easy using the Axiom of Specification to create the empty set. Then he introduces the Axiom of Pairing, which states that if we have two sets, there exists a set containing both. Having created such a set, he names it an unordered pair. He points out than any unordered pair where both elements are the same (Ex. $\{a, a\}$) is a singleton, and then uses this to construct the familiar series of sets of variations off the empty set, such as $\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}$.

He closes with a point on notation, introducing the $\{x : S(x)\}$ notation.

3.4 Chapter 4: Unions and Intersections

Halmos starts off by introducing the Axiom of Unions, $\forall \mathcal{C}, \exists Y = \{y : \exists X \in \mathcal{C}, y \in X\}$. He then proves some obvious truths about this operation, such as

$\cup\emptyset = \emptyset, \cup\{A\} = A, \cup\{A, B\} = A \cup B, A \cup \emptyset = A$.He also states that the $\cup$ 'relation' has commutativity, associativity, idempotence, and related to the $\subset$ relation in the following way:

Theorem 3.3 *$A \subset B \iff A \cup B = B$.*

Theorem 3.4 *Starting with the $\Rightarrow$ direction, we see trivially that $B \subset A \cup B$. Then since $A \subset B \rightarrow (a \in A \implies a \in B)$, $c \in A \cup B \rightarrow c \in A \vee c \in B \rightarrow c \in B \vee c \in B \rightarrow c \in B$, therefore $A \cup B \subset B$ and so $A \cup B = B$.*

Now the $\Leftarrow$ direction. Assuming $A \cup B = B$, that means $c \in A \cup B \implies c \in B$. Then $c \in A \implies c \in A \cup B \implies c \in B$, therefore $A \subset B$.

He points out that using this $\cup$ operation we can construct triples, quadruples, and so on. He also develops the $\cap$ relationship in a unique way. Where I've usually seen it given as $A \cap B = \{x \in A \cup B : x \in A \wedge x \in B\}$, he defines it as $A \cap B = \{x \in A : x \in B\} = \{x \in B : x \in A\}$. Of course, both are identical and neither one is more robust than the other. He then states the same properties for $\cap$ as he did for $\cup$.

He then defines the distributivity laws and the property of disjointedness (A and B disjoint if $A \cap B = \emptyset$). He then defines the broader intersection operation, $\cap\mathcal{C}$, such that $\cap\mathcal{C} = \{x : \forall X \in \mathcal{C}, x \in X\}$.

3.5 Chapter 5: Complements and Powers

Halmos begins the chapter discussing set differences and the relative complement. He emphasizes this relative complement is always taken in reference to an overarching set, ie. there is no A', only A' with respect to E. Here A' is defined as $\{x \in E : x \notin A\}$. He then gives some obvious truisms about the complement and also states De Morgan's Laws. He also defines the boolean sum of $A + B = (A - B) \cup (B - A)$. He states the usual properties of this operation, associativity, commutativity, and the identity relationship $A + \emptyset = A$.

Then he introduces the Axiom of Powers: $\forall E, \exists P = \{X : X \subset E\}$. He gives the usual terminology and basic facts for this set P, P is the 'power set of E', and if E has n elements, P has 2^n elements.

Theorem 3.5 *$\mathcal{P}(E) \cap \mathcal{P}(F) = \mathcal{P}(E \cap F)$.*

Take $X \in \mathcal{P}(E) \cap \mathcal{P}(F)$, then $X \subset E$ and $X \subset F$, so $x \in X \Rightarrow x \in E$ and $x \in F \Rightarrow x \in E \cap F \Rightarrow X \subset E \cap F \Rightarrow X \in \mathcal{P}(E \cap F) \Rightarrow \mathcal{P}(E) \cap \mathcal{P}(F) \subset \mathcal{P}(E \cap F)$.

Going the other direction, take $X \in \mathcal{P}(E \cap F)$, then $X \subset E \cap F \Rightarrow \forall x \in X, x \in E \wedge x \in F \Rightarrow X \subset E \wedge X \subset F \Rightarrow X \in \mathcal{P}(E) \wedge X \in \mathcal{P}(F) \Rightarrow X \in \mathcal{P}(E) \cap \mathcal{P}(F) \Rightarrow \mathcal{P}(E \cap F) \subset \mathcal{P}(E) \cap \mathcal{P}(F)$.

$\Rightarrow \mathcal{P}(E) \cap \mathcal{P}(F) = \mathcal{P}(E \cap F)$

3.6 Chapter 6: Ordered Pairs

Halmos now defines ordered pairs in the standard way: if we want an ordered set (a, b, c) we define it as $\{\{a\}, \{a, b\}, \{a, b, c\}\}$.

Theorem 3.6 $(a, b) = (c, d) \iff a = c \land b = d$.

$\Leftarrow$ is trivial. To prove $\Rightarrow$, observe that $(a, b) = (c, d)$ means $\{\{a\}, \{a, b\}\} = \{\{c\}, \{c, d\}\}$, so either $\{a\} = \{c\}$ or $\{a\} = \{c, d\}$, therefore either way $a = c$ (it's possible that $a = d$ as well). Then $\{a, b\} = \{c\}$ or $\{c, d\}$.

If $\{a, b\} = \{c\}$, then $a = b = c$ and $(a, b) = \{\{a\}\}$. Since the pairs are equal, we also know that $\{c, d\} = \{a\}$ and so $c = d = a$. So we've shown $a = c \land b = d$.

If $b \neq a$ and $\{a, b\} = \{c, d\}$, then $a = c \to b = d$. So in either case we have shown the $\Rightarrow$ holds.

Halmos also shows the existence of the Cartesian product by the fact that for $\forall a \in A, \{a\} \subset \mathcal{P}(A)$.

Finally, he also defends the practice constructing the concept of an ordered pair instead of taking it as axiomatic. While this does reduce the number of needed axioms, he points out that there are a few "pathological" features of the construction, such as the fact that $\{a, b\} \in (a, b)$.

3.7 Chapter 7: Relations

Here Halmos introduces the idea of a binary relation. Specifically here he constructs a relation as a set of ordered pairs. If R is a relation and $(x, y) \in R$, then x 'has that relation' to y. He goes on to define the domain and range, such that $dom(R) = \{x : y, (x, y) \in R\}$ and $ran(R) = \{y : x, (x, y) \in R\}$. He extends the usual definitions of reflexivity, transitivty, and symmetry to relations, and defines an equivalence relation as one which has all three properties.

He then introduces the idea of a partition, which is a split of a set X in nonempty, disjoint subsets (C_α) such that $\bigcup_{\alpha \in I} C_\alpha = X$. This then interacts with the concept of an equivalence class, which is defined as (for $x \in X$, and a relation R) $x/R = \{y : y \in X, (x, y) \in R\}$.

Theorem 3.7 *For any set A and any equivalence relation R, R induces a partition of A, and any partition P of A can be used to define an equivalence relation* R_P.

1. Take any element a of A, then since $a\text{R}a$, a is in at least one equivalence class, the goal is to show that a is in at most one equivalence class. Assume $a \in X \land a \in Y$, where both X, Y are equivalence classes and $X \neq Y$. Then $\exists x$ W.L.O.G $x \in X, x \notin Y$ such that xRa, however, since an equivalence relation must be transitive, $\forall y \in Y, xRa \land aRy \to xRy \to y \in X$. Therefore $X \subset Y$, which contradicts our assumption. Therefore a is in exactly one equivalence class, and since it was chosen arbitrarily, this shows that R induces a partition of A.

2. Take a partition P of A. Then define R_P as $\{(x, y) : \exists X \in P, x \in Y \land y \in Y\}$. Clearly this reflexive, since every element must be in one set of the partition. It is also transitive, since $aRb \land bRc$ implies a and c are both in the group containing b and therefore aRc. Finally, it is symmetric, since if aRb, then they both are in the same set of the partition, and so bRa. Thus we have defined an equivalence relation based on the partition.

3.8 Chapter 8: Functions

Halmos opens by defining a function from $X \to Y$ as a relationship with domain X and range Y, and such that each $x \in X$ only relates to a single $y \in Y$. Having defined this concept, he brings over the familiar domain and range definitions from his discussion of relations. He then brings in 'image of', notation, such as $f(X) = Y$, but addresses that the standard notation is fairly bad (his example is, if $A \subset X$ AND $A \in X$, what does $f(A)$ refer to?).

He then introduces inclusion maps, defined as $X \subset Y, f : X \to Y, f(x) = x$. He identifies the identity map $f : X \to X, f(x) = x$ as a special case of inclusion maps. He also defines function extensions and restrictions in the usual way, g is a restriction of f if $A \subset X, f : X \to Y, g : A \to Y, g(x) = f(x)$. He also defines projections, which is a function from $f : X \times Y \to X$ s.t. $f(x, y) = x$. He also defines canonical maps for X and an equivalence relation on X, R. Then $f : X \to X/R, f(x) = x/R$ is the canonical map. The canonical maps are important because for a function f, defining an equivalence relation R s.t. $x, y \in X, (x, y) \in R \iff f(x) = f(y)$ gives us a partition of X based on where each element is mapped to by f.

Theorem 3.8 *Take $f : X \to Y$. The canonical map $g : Y \to X/R, g(y) = \{x \in X : f(x) = y\}$ is one-to-one.*

Take $x, y \in Y, g(x) = g(y)$. Then take any element $a \in g(x)$, then by how we defined g, $f(a) = x \wedge f(a) = y \Rightarrow x = y$. Thus g is bijective.

He then introduces the characteristic function, and uses the notion of canonical mappings to prove there's a one-to-one relationship between a powerset $\mathcal{P}(X)$ and the set 2^X, which will surely be useful later.

3.9 Chapter 9: Families

Here Halmos defines the idea of a family, which is a mapping from an index set I to another set X. Specifically, he introduces the terminology of a family of elements $\{x_i\}_{i \in I}$ in X or a family of subsets $\{A_i\}$ which we can take the union of: $\bigcup_{i \in I} A_i$. He goes on to show that talking about collections of sets or families are equivalent, since any collection of sets is also the family of sets defined by some index set.

Theorem 3.9 *$B \cap \bigcup_i A_i = \bigcup_i (B \cap A_i)$ and $B \cup \bigcap_i A_i = \bigcap_i (B \cup A_i)$.*

1. Take $a \in B \cap \bigcup_i A_i$. Then $a \in B$ and $\exists k : a \in A_i$. Therefore $a \in B \cap A_i \Rightarrow a \in \bigcup_i (B \cap A_i) \Rightarrow B \cap \bigcup_i A_i \subset \bigcup_i (B \cap A_i)$. Then take $a \in \bigcup_o (B \cap A_i)$, so $\exists k : a \in (B \cap A_i) \Rightarrow a \in B \wedge a \in A_i \Rightarrow a \in \bigcup_i A_i \Rightarrow a \in B \cap \bigcup_i A_i \Rightarrow \bigcup_i (B \cap A_i) \subset a \in B \cap \bigcup_i A_i$.

2. Take $a \in B \cup \bigcap_i A_i$, then $a \in B \vee \forall i, a \in A_i, \Rightarrow \forall i, a \in (B \cup A_i) \Rightarrow a \in \bigcap_i (B \cup A_i)$ and so $B \cup \bigcap_i A_i \subset \bigcap_i (B \cup A_i)$. Now take $a \in \bigcap_i (B \cup A_i) \Rightarrow \forall i, a \in B \vee a \in A_i$. Therefore $a \in B \cup \bigcap_i A_i \Rightarrow \bigcap_i (B \cup A_i) \subset B \cup \bigcap_i A_i$.

Halmos also uses families to generalize a Cartesian product, which is used to create the usual sense of 'coordinates' for a function of several variables, which can be thought of as a function with a domain of a Cartesian product of a family.

3.10 Chapter 10: Inverses and Composites

Here Halmos introduces the idea of an inverse mapping specifically for $f : X \to Y$, the inverse is $f^{-1} : \mathcal{P}(Y) \to \mathcal{P}(X)$ such that for $y \subset Y$, $f^{-1}(y) = x \in X : f(x) \in y$. He calls this $f^{-1}(y)$ the inverse image of y under f. If f is one-to-one, he allows a second definition where $f^{-1} : Y \to X$ such that $f^{-1}(y) = x : f(x) = y$.

Having defined the inverse, he now defines a composite function. If $f : X \to Y, g : Y \to Z$, then the composite function $g(f(x))$ is denoted $gf : X \to Z$.

Theorem 3.10 *The inverse of a composite function is the composite of the component functions' inverses, ie.* $(gf)^{-1} = f^{-1}g^{-1}$.

Take invertible functions $g, f,\ f : A \to B, g : B \to C$. Then take their composition $gf : A \to C$. Observe that $\forall x \in C$, $(f^{-1}g^{-1})(x) = f^{-1}(g^{-1}(x))$ is equal to some $y \in A$. Then $(gf)(y) = g(f(y)) = g(f(f^{-1}(g^{-1}(x))))$. But notice that by definition of the inverse, $\forall z, f(f^{-1}(z)) = z$ and $g(g^{-1}(z)) = z$. Therefore $g(f(f^{-1}(g^{-1}(x)))) = g(g^{-1}(x)) = x$. And therefore $f^{-1}g^{-1}$ maps elements of C precisely onto their inverse image under gf in A, and therefore it is the inverse.

He generalizes this to the idea of relations. The idea that for a relation R, $(y, x) \in R^{-1} \iff (x, y) \in R$. He shows this 'switches' the domain and range of R. He then proves that the properties of inverse and composite functions also extend to inverse and composite relationships.

3.11 Chapter 11: Numbers

In this chapter, Halmos finally begins to define the concept of a number. He introduces the successor function, where $x^+ = x \cup \{x\}$. Then, by setting $0 = \emptyset$, successive iterations of the successor function create the usual definitions of $1 = \{\emptyset\} = \{0\}, 2 = \{0, 1\}, 3 = \{0, 1, 2\}$ and so on.

However, here Halmos introduces another of the ZF axioms, the axiom of infinity, which Halmos formulates as "There exists a set containing 0 and containing the successor of each of its elements" (43). Once we take the intersection of all such sets, it essentially means we can speak of N as a set, and assume it exists, whereas otherwise we could only say that a set containing all the successors of N up to an arbitrary (but not infinity) number of iterations.

He ends by defining a sequence as a family whose index set is a natural number or the set of all natural numbers, which is the standard definition.

3.12 Chapter 12: The Peano Axioms

First Halmos defines the natural numbers ω as (1) a set which contains 0, and (2) $x \in \omega \rightarrow x^+ \in \omega$, and finally (3) if $S \subset \omega, [0 \in S \wedge (n \in S \rightarrow n^+ \in S)] \rightarrow S = \omega$. He describes this last property (which is equivalent to showing ω is the minimal successor set) as the "principal of mathematical induction" (46).

He then proves two additional properties of ω. First is that (4) if $n \in \omega, n^+ \neq 0$, which is simply shown since $n \in n^+$ and $n \notin 0 = \emptyset$.

Lemma 3.11 *If $n \in \omega$ and $m \in n$, then $n \not\subset m$.*

This is vacuously true for 0, and then if it holds for n, it must hold for n^+, since otherwise if $\exists m \in n^+$ s.t. $n^+ \subset m$, $m = n$ or some $x \in n$. If $n^+ \subset x, thenn \subset x \in n$ which is a contradiction. Similarly $n^+ \subset n$ would imply $n \in n$, which contradicts our induction hypothesis since $n \subset n$. Therefore, by our induction property, the set of natural numbers that have this property consists of all of ω.

Lemma 3.12 *If $n \in \omega, m \in n \rightarrow m \subset n$.*

We see again this is vacuously true for 0. Then if this is true for n, any $m \in n^+$ is either $\in n$ and therefore a subset of n and n^+, or is equal to n and therefore is still a subset of n^+. Therefore by induction this property holds for all natural numbers.

Theorem 3.13 *If m, n are natural numbers and $n^+ = m^+$, then $n = m$.*

Assume this is untrue, which would imply $n \in m$ and $m \in n$, which then since by the second lemma $m \subset n$, $n \in n$ and $n \subset n$ but this contradicts the first lemma, so we must have $n = m$ as desired.

These five properties combined composed the Peano Axioms (somewhat of a misnomer here, since we proved them). Using these, Halmos establishes another theorem, the Recursion theorem, which states that if $f : X \rightarrow X$ and $a \in X$, then $\exists u : \omega \rightarrow X, u(n^+) = f(u(n))$. Essentially this allows induction to be used constructively to define objects.

3.13 Chapter 13: Arithmetic

Using this Recursion Theorem, Halmos goes on to describe the typical properties of arithmetic on the natural numbers. He starts with addition, defined as a function for every $m \in \omega$ s.t $s_m(0) = m, s_m(n^+) = (s_m(n))^+$, and shows it is commutative by a simple induction proof. Similarly he defines multiplication as $p_m(0) = 0$ and $p_m(n^+) = p_m(n) + m$. We are assured without proof that the expected properties and definitions for these operations hold.

He also establishes that natural numbers are always comparable, ie. $m, n \in \omega \rightarrow (m \in n) \vee (n \in m) \vee (n = m)$. Using these Halmos defines the notation for $n < m \iff n \in m, m < n \iff m \in n$ and defined $\leq$ and $\geq$ in their usual ways.

Halmos defines equivalence between sets as the existence of a one-to-one mapping between them. Then he states that any subset of a natural number is equivalent to a smaller natural number, which is proved easily by induction. While he points out sets can be equivalent to proper subsets of themselves, such as mapping the natural numbers to the non-zero natural numbers, he states that a finite set (ie. a set equivalent to a natural number) cannot be equivalent to a proper subset of itself.

Theorem 3.14 *A set can be equivalent to at most one natural number.*

We will first show that equivalence is (appropriately) an equivalence relation. Take a set A, then the mapping $f : A \to A, f(a) = a$ is clearly one-to-one and so $A \sim A$. Every one-to-one mapping is invertible, so if $A \sim B$, there exists $f : A \to B$, f one-to-one, and then $f^{-1} : B \to A$ is a one-to-one mapping so $B \sim A$. Finally, if $A \sim B \wedge B \sim C$. Then $\exists f : A \to B$ and $g : B \to C$, both one-to-one. Take $gf : A \to C$. Then if $g(f(x)) = g(f(y))$, g being one-to-one implies $f(x) = f(y)$, and f being one-to-one implies $x = y$. Therefore gf is one-to-one and so $A \sim C$. Therefore equivalence is an equivalence relation, and therefore can be used to induce a partition on ω. However, since each natural number is comparable under proper subset ordering, and a finite set cannot be equivalent to a proper subset of itself, the partition of ω by equivalence is only a set of singletons. This then implies that a set cannot be equivalent to two different elements of ω, since otherwise by the transitivity, there would be two elements of ω in the same equivalence class.

Halmos also asserts that a subset of a finite set is finite.

3.14 Chapter 14: Order

Here Halmos introduces the concept of an order, based on the previous definitions of relations. He defines an antisymmetric relationship as one where $xRy \wedge yRx \iff x = y$. Then he states a partial order is a relationship on a set X such that the relationship is antisymmetric, reflexive, and transitive. If every element can be compared to every other element under this relation, then it is a total order.

He then defines sets with a partial order as a partially ordered set. He then demonstrates that for any ordering that includes $\leq$ relations, we can define a relationship that includes all the same relations except when the elements are equal, thus creating a strict relationship. This can also be done in reverse.

He defines an initial segment of a in a partially ordered set S as $\{x \in S : x < a\}$. He also states that if $a \leq x, \forall x \in S$, we can refer to a as a least element of S, and likewise we can define the concept of a greatest element. He distinguishes between a minimal element and a least element, specifically pointing out that while there can only be one least element, there can be many minimal elements.

He goes on to define upper and lower bounds, and then defines the supremum and infimum as the least and greatest upper and lower bounds respectively.

3.15 Chapter 15: The Axiom of Choice

Halmos formulates the Axiom in a Choice in a somewhat unique way, as "the Cartesian product of a non-empty family of non-empty sets is non-empty" (59). This is of course equivalent to the traditional axiom of choice, in that if we can take a collection $\mathcal{C}$, then by taking the Cartesian product of the family of the collection indexed by itself, then any element in the Cartesian product represents a 'choice' of an element from each set in the collection.

He then demonstrates the proper use of the Axiom of Choice in proofs by completing a sample proof of the fact that any infinite set has a subset equivalent to ω.

3.16 Chapter 16: Zorn's Lemma

Halmos opens here by with the statement of Zorn's Lemma.

Theorem 3.15 (Zorn's Lemma) *If X is a partially ordered set such that every chain in X has an upper bound, then X contains a maximal element.*

First begin by taking the initial segment $s(x)$ of each $x \in X$. This can be thought of as a function $s : X \to \mathcal{P}(X)$. Since this is a one-to-one function and $s(x) \subset s(y) \iff x \leq y$, a maximal element in X corresponds to a maximal element in $s(X)$ when partially ordered by inclusion. Then take the set of all chains in X, denoted $\mathcal{X}$.

Take a choice function on $\mathcal{X}$. Then denote $\bar{A}$ as the set $\{x \in X : A \cup \{x\} \in \mathcal{X}\}$. Now define a function $g : \mathcal{X} \to \mathcal{X}$, such that $g(A) = A \cup \{f(\bar{A} - A\}$. Then clearly if $g(A) = A$, A must be a maximal element.

Next we develop the idea of a tower, as an element of $\mathcal{X}$ such that a tower must contain $\emptyset$ and contains $g(A)$ if it contains A, and also contains the union of all subset of any chain it contains.

Take the intersection of all towers $\mathcal{T}_0$. Then take an element $C \in \mathcal{T}_0$ such that C is comparable to any element of $\mathcal{T}_0$. Then take an element $A \in \mathcal{T}_0, A \in C$. Then $g(A) \subset C$. Take $\mathcal{U} = \{A \in \mathcal{T}_0 : A \subset C \vee g(C)A\}$. Then $\mathcal{U}$ is a tower and $\mathcal{T}_0 \subset \mathcal{U}$, but since $\mathcal{T}_0$ is the smallest tower, $\mathcal{T}_0 = \mathcal{U}$. This then demonstrates that $g(C)$ must also be comparable to any element of $\mathcal{T}_0$.

Thus the comparable sets in $\mathcal{T}_0$ exhaust $\mathcal{T}_0$ and $\mathcal{T}_0$ is itself a chain. Therefore if A is the union of all sets in $\mathcal{T}_0$, then it's included in $\mathcal{T}_0$, but it also contains every set in $\mathcal{T}_0$, so $g(A) \in A$ and $A \in g(A)$, so A is a maximal element as desired.

3.17 Chapter 17: Well Ordering

Here Halmos defines well ordering as the property that any non-empty subset has a least element. He then asserts the conditions under which transfinite induction is possible, which is primarily that there exists X well-ordered, and $S \subset X$ such that $\forall x \in X : s(x) \subset S \to x \in S$, where $s(x)$ is the initial segment of x. If these conditions are met, then $S = X$.

Halmos makes a concerted effort here to emphasize that this is not equivalent to regular induction. While the above conditions do on ω allow for the same conclusions, simply applying regular induction to a well-ordered set does not allow for transfinite induction.

He proves transfinite induction quite easily, since the proof is just that if $X - S$ is non-empty, it has a least element x, then of course $s(x) \in S$, therefore $x \in S$. This is a contradiction however, so $X - S$ is empty and therefore $S = X$.

He also proves that the union of a chain of well-ordered sets must be well-ordered, although the 'chain' here refers to an ordering based on a concept called continuation. A well-ordered set A is a continuation of another B if $B \subset A$, $\exists a \in A : s(a) = B$, and the ordering of B is the same as in A.

Halmos then gets into the eponymous theorem of the chapter. The proof depends explicitly on the Zorn Lemma and therefore the Axiom of Choice.

Theorem 3.16 (Well-Ordering) *Every set can be well-ordered.*

For a set X take $\mathcal{W}$ such that $\mathcal{W}$ is the set of all well-ordered subset of X. Then order $\mathcal{W}$ by continuation. If $\mathcal{C}$ is a chain, then $\cup\mathcal{C} \geq C, \forall C \in \mathcal{C}$. Thus $\cup\mathcal{C}$ is a maximal element for $\mathcal{C}$, and so Zorn's lemma applies and there is a maximal element M of $\mathcal{W}$. This M must be equal to the entire set or else there is some element $x \in X, x \notin M$, and by placing x directly after each element of M we have created a larger element of $\mathcal{W}$, which would contradict that M is maximal. Therefore $M = X$ and $M \in mathcalW$ implies X can be well-ordered.

3.18 Chapter 18: Transfinite Recursion

Halmos begins by defining the Transfinite Recursion Theorem: "If W is a well ordered set, and if f is a sequence function of type W in a set X, then there exists a unique function U from W into X such that $U(a) = f(U^a)$ for each a in W."

3.19 Chapter 19: Ordinal Numbers

Halmos begins by asking what would happen if we took the successor function of the entire set of natural numbers, ie. ω^+, and whether there is a set which contains all such successors of the natural numbers. Halmos defines a ω-successor function f as $f(0) = \omega$, $\text{dom} f = n \in \omega$ and $\forall m : m < n, f(m^+) = (f(m))^+$.

Theorem 3.17 *There exists an ω-successor function for any natural number.*

Let S be the set of natural numbers s.t.there exists a unique ω-successor function. Then the statement vacuously holds that $0, 1 \in S$. Now assume that $n \in S$, and we will show $n^+ \in S$.

Let $\text{dom}(f_{n^+}) = n^+$, and for $m \in n^+$, $f_{n^+}(m) = f_n(m)$. Then define $f(n^+) = f(n) \cup \{f(n)\} = (f(n))^+$. Then this is a unique ω-successor function, since f_n is unique and if $f_{n^+}(n^+) \neq (f(n))^+$ then f_{n^+} is not an ω-successor function. Therefore $S = \omega$. So there exists an ω-successor function for any natural number.

To make use of this, Halmos introduces the Axiom of Substitution, which states that for a sentence $S(a, b)$ such that $\{b : b \in A \wedge S(a, b)\}$ exists, there exists F with domain A such that $F(a) = \{b : S(a, b)\}$. Using this Axiom and the above ω-successor functions, the existence of a function F with domain ω that maps each natural number to $\omega + n$ follows (the first infinite ordinals). Halmos calls the range of this function $\omega 2$.

3.20 Chapter 20: Sets of Ordinal Numbers

Halmos begins examining the properties of an ordinal number. The most elementary property is that an ordinal number must be a transitive set, and is actually the set of its predecessors under the ordering of the ordinal number. This itself means that each element of an ordinal number is itself an ordinal number.

Theorem 3.18 *Similarity between ordinal numbers implies equality.*

Take a similarity f from $\alpha \to \beta$. Then $\forall \xi \in \alpha$, the least element of $\alpha \notin s(\xi) = \xi$ and the least element of $f(\alpha)$, not in $f(s(\xi))$ must be $f(\xi)$. Then ξ and $f(\xi)$ must both be ordinal numbers with the same initial segments and $\xi = f(\xi)$. Letting $S = \{\xi \in \alpha : f(\xi) = \xi\}$, by transfinite induction we have shown that $S = \alpha$ and therefore $\alpha = \beta$.

Furthermore being ordinals for any two ordinals α and β are either similar or one is similar to an initial of the other. This means all ordinal numbers are comparable.

Then he establishes the counting theorem, which states that every well-ordered set is similar to an ordinal number.

3.21 Chapter 21: Ordinal Arithmetic

By defining $\widehat{E} = \{(e, 0) : e \in E\}$ and $\widehat{F} = \{(f, 1) : f \in F\}$, and two sets can be made disjoint while still preserving their order and other important properties. Using this we can the define an addition between them as $\widehat{E} \cup \widehat{F}$ with the same order structure and the property that every element of $\widehat{E}$ is less than $\widehat{F}$, then clearly this set is also well-ordered and more than that, it is transitive and an ordinal number.

Specifically, we define the sum of $E + F$ as the ordinal number which is similar to $\widehat{E} \cup F$. There are a couple worthwhile properties to show, whose (somewhat trivial) proofs Halmos omits:

1. $\alpha \cup \emptyset = \alpha \Rightarrow \alpha + 0 = \alpha$
2. $\emptyset \cup \alpha = \alpha \Rightarrow 0 + \alpha = \alpha$
3. $\alpha \cup \{\emptyset\} = \alpha \cup \{\beta\}$ where $\beta > \alpha$ and $\alpha \in \alpha \cup \{\beta\} \Rightarrow \beta = \alpha$. Therefore $\alpha + 1 = \alpha^+$.

However, he points out there are some less nice properties of this addition system, such as the lack of commutativity ($1 + \omega = \omega$, even as $\omega + 1 = \omega^+$ as we showed above).

From this we can define the ordinal product $A * B$ as $\sum_{b \in B} A_b$ where $A_b = \{(a, b) : a \in A\}$. Halmos points out this is equivalent to $A * B$ being defined as $\{A \times \{b\} : b \in B\}$ with the reverse lexicographical order. He gives similar multiplicative identities as we had for addition:

1. $A * 0 = 0$
2. $0 * A = 0$
3. $A * 1 = A$
4. $1 * A = A$
5. $A(B\gamma) = (AB)\gamma$
6. $A(B + \gamma) = AB + A\gamma$

However, like multiplication, commutativity fails: $2\omega = \omega \neq \omega 2$. And the right distributive law also fails, the example Halmos gives being $(1 + 1)\omega = (2)\omega = \omega \neq \omega 2 = 1\omega + 1\omega$. By the same recursive technique Halmos defines exponentiation, and gives the following identities without proof:

1. $0^\alpha = 0$
2. $1^\gamma = 1$
3. $\alpha^{\beta+\gamma} = \alpha^\beta + \alpha^\gamma$
4. $\alpha^{\beta\gamma} = (\alpha^\beta)^\gamma$

However, once again the operation does not fulfill every standard identity. For example Halmos gives the fact that $(2*2)^\omega = 4^\omega = \omega \neq 2^\omega * 2^\omega = \omega * \omega = \omega^2$.

3.22 Chapter 22: The Schröder-Bernstein Theorem

Halmos begins by describing why ordinal numbers aren't exactly useful for measuring the 'number' of elements in a set. He gives the example of ω ordered by placing 0 after every other element. This new ordering has ordinal number $\omega+1$, even though it intuitively has the same 'number' of elements. He blames this on the concept that the ordinal number of a well-ordered set is determined by what it is similar to (specifically whether or not it is similar to an initial segment of another well-ordered set).

Halmos defines a relation between X and Y $X \preceq Y$ if X is equivalent to a subset of Y (equivalent, as previously defined, meaning there is a one-to-one mapping between them). The Schrödinger-Bernstein Theorem develops a relationship between this relation, called domination, and equivalence of the larger sets. It states that.

Theorem 3.19

$$X \preceq Y \wedge Y \preceq X \Rightarrow X \sim Y$$

Take one-to-one into mappings $f : X \to Y, g : Y \to X$. For $x \in X$, call x the 'ancestor of any element y in X, Y if y is the result of alternatively applying f and g to x an arbitrary amount of times. We also say y is a descendant of x. For any element we can trace its 'ancestry' back until we either reach an element in $X/f(Y), Y/f(X)$, or the ancestry chain will go infinitely. Partition X and Y into these categories, X_X, X_Y, X_∞ and likewise for Y.

Then for $x \in X_X$, $f(x) \in Y_X$ since they share an ancestry chain, and so $f|_{X_X}$ is a one-to-one mapping between X_X and Y_X. $x \in X_Y$, then there must exist $y \in Y_Y$ such that $g(y) = x$ and so $g^-1|_{X_Y}$ is a one-to-one mapping between X_Y and Y_Y. Then finally by the same logic f is a one-to-one mapping between X_∞ and Y_∞, and combining these three mappings together we obtain a one-to-one mapping between X and Y so $X \sim Y$.

3.23 Chapter 23: Countable Sets

Halmos begins with a formal description of countability. A set X is countable if $X \preceq \omega$ or $X \sim \omega$. He shows a simple proof that any subset of a countable set is also countable. He also shows the union of countably infinite family of countable sets is countable.

He then introduces Cantor's Theorem.

Theorem 3.20 (Cantor's Theorem) $\forall X, X \prec \mathcal{P}(X)$.

Assume there exists an f from X onto $\mathcal{P}(X)$. Let $A = \{x \in X, x \notin f(x)\}$. Then since f is onto there exists $a \in X$ such that $f(a) = A$. Then this is a contradiction since $a \in A \implies a \notin f(a) = A$ and $a \notin A \implies a \in f(a) = A$. Therefore there must not exist such a function.

Since we can map $x \in X \to \{x\} \in \mathcal{P}(X)$, we know that $\mathcal{P}(X)$ dominates X, therefore the theorem is proved.

3.24 Chapter 24: Cardinal Arithmetic

Here Halmos promises a construction of Cardinal numbers which allows an ordering, and has the property that card $X =$ card $Y \iff X \sim Y$, although he postpones the construction for a later chapter.

Instead, he defines cardinal addition in the same way he did for ordinal addition. Specifically, if $\alpha = \text{card}A$ and $\beta = \text{card}B$, then $\alpha + \beta = \text{card}(A \cup B)$.

His definition for cardinal multiplication however is different from ordinal multiplication. If we take α and β as defined in the paragraph above, then $\alpha\beta = \text{card}(\alpha \times \beta)$. Finally, he defines exponents as $\alpha^\beta = \text{card}A^B$.

3.25 Chapter 25: Cardinal Numbers

Here Halmos begins his actual definition of the cardinal numbers. He points out that a set can be equivalent to many different ordinal numbers (see the ω example above). However, by transitivity of equivalence, and the Cantor Theorem, we know that every ordinal number which is equivalent to A is strictly dominated by $\mathcal{P}(A)$. Then since $\mathcal{P}(A) \sim \gamma$ for some ordinal γ, we know every ordinal equivalent to A is an element of γ. Then by the axiom of separation we can construct a set from γ of each ordinal which is equivalent to A.

We then call the least element of this set (which exists since the ordinals are well-ordered) the cardinal number of A.

He finally ends by discussing the notation for cardinals, which are usually denoted by $\aleph$, with $\omega = \aleph_0$. He ends with the continuum hypothesis, offering no support in favor or against it- the hypothesis states that $\aleph_1$ (the smallest uncountable cardinality) is equal to 2^ω.

4 Application of Transfinite Induction to the Borel Hierarchy

4.1 Preliminaries

We define a σ-algebra as a subset which contains the set itself, and is closed under complement and countable unions.(3) Define $\mathcal{B}(R^n)$ as the smallest σ-algebra of R^n that contains the open sets. Call a member of this σ-algebra a *Borel Set*.

We define the Borel Hierarchy in R^n as follows (4):

1. Take $\sum_{\mathbf{1}}^{\mathbf{0}}$ as the collection of open sets in R^n.

2. Take $\prod_{\boldsymbol{\alpha}}^{\mathbf{0}}$ as the collection of complements in the corresponding $\sum_{\boldsymbol{\alpha}}^{\mathbf{0}}$.

3. Let $\sum_{\boldsymbol{\alpha}}^{\mathbf{0}}$ be the collection of sets that can be constructed $\bigcup_i A_i$, where each $A_i \in \prod_{\boldsymbol{\alpha_i}}^{\mathbf{0}}$ for some $\alpha_i < \alpha$.

4. Let $\boldsymbol{\Delta}_{\boldsymbol{\alpha}}^{\mathbf{0}} = \sum_{\alpha}^{0} \cup \prod_{\boldsymbol{\alpha}}^{\mathbf{0}}$.

We let α grow to any ordinal less than or equal to ω_1, and we seek to prove that $\sum_{\boldsymbol{\omega_1}}^{\mathbf{0}} = \prod_{\boldsymbol{\omega_1}}^{\mathbf{0}} = \mathcal{B}$. We seek to prove two facts about this construction:

1. Every set in the Hierarchy is a member of this sigma algebra. $\iff \sum_{\boldsymbol{\omega_1}}^{\mathbf{0}} \subset \mathcal{B}$

2. Every member of this sigma algebra is in the Hierarchy. $\iff \mathcal{B} \subset \sum_{\boldsymbol{\omega_1}}^{\mathbf{0}}$

4.2 All Sets in the Hierarchy are Borel Sets

Theorem 4.1 *All sets in an ordinal rank of the Borel Hierarchy are Borel Sets*

We observe trivially that every set in $\sum_{\mathbf{1}}^{\mathbf{0}}$ and $\prod_{\mathbf{1}}^{\mathbf{0}}$ must be Borel sets since they are respectively the open and closed sets in R^n.

Furthermore, if $\sum_{\boldsymbol{\alpha}}^{\mathbf{0}}$ is composed of only Borel Sets, $\prod_{\boldsymbol{\alpha}}^{\mathbf{0}}$ must be composed of only complements of Borel Sets and therefore every set in it is Borel, and therefore $\sum_{\boldsymbol{\alpha+1}}^{\mathbf{0}}$ must also be made up of Borel Sets, since it is composed of only countable unions of Borel Sets.

Finally, observe that if we have a limit ordinal α, and we know $\sum_{\boldsymbol{\beta}}^{\mathbf{0}}$ and $\prod_{\boldsymbol{\beta}}^{\mathbf{0}}$ contain only Borel sets for $\beta < \alpha$, then $\sum_{\boldsymbol{\alpha}}^{\mathbf{0}}$ consists of only countable unions of Borel sets, and therefore is itself composed of Borel sets, and therefore likewise for $\prod_{\boldsymbol{\alpha}}^{\mathbf{0}}$.

Thus we have shown that this property holds for $\sum_{\mathbf{1}}^{\mathbf{0}}$ and $\prod_{\mathbf{1}}^{\mathbf{0}}$, that if it holds for an ordinal α, it must hold for $\sum_{\boldsymbol{\alpha+1}}^{\mathbf{0}}$ and $\prod_{\boldsymbol{\alpha+1}}^{\mathbf{0}}$, and that if it holds for all ordinals less than a limit ordinal, it must hold for that ordinal. Thus by transfinite induction, this property holds for all ordinals.

4.3 The Borel Hierarchy Contains all Borel Sets

The basic structure of our proof was found in (5). Here we simplify several elements as well as prove several assumptions the text makes. First we establish several lemmas.

Lemma 4.2 *Any closed set is countable intersection of open sets*

Take an arbitrary closed set C. Define $O_n = \{x \in R^n : \inf_{c \in C} d(x, c) < \frac{1}{n}\}$. Then if $x \in \bigcap O_n$, $\inf_{c \in C} d(x, c) = 0$, and so $\forall n, \exists c_n$ s.t. $d(c_n, x) < \frac{1}{n}$, but then x is a limit point of C, and since C is closed, $x \in C$. Therefore $\bigcap O_n \subset C$, and since $C \subset \bigcap O_n$ trivially, $\bigcap O_n = C$ and since C is arbitrary, we have shown any closed set is a countable intersection of open sets.

Lemma 4.3 *For two ordinals α and β, $\alpha < \beta$, $\sum_{\boldsymbol{\alpha}}^{\mathbf{0}} \subset \sum_{\boldsymbol{\beta}}^{\mathbf{0}}$, $\sum_{\boldsymbol{\alpha}}^{\mathbf{0}} \subset \prod_{\boldsymbol{\beta}}^{\mathbf{0}}$, $\prod_{\boldsymbol{\alpha}}^{\mathbf{0}} \subset \prod_{\boldsymbol{\beta}}^{\mathbf{0}}$, and $\prod_{\boldsymbol{\alpha}}^{\mathbf{0}} \subset \sum_{\boldsymbol{\beta}}^{\mathbf{0}}$*

Observe that every closed set is in $\prod_{\mathbf{1}}^{\mathbf{0}}$, since every open set is in $\sum_{\mathbf{1}}^{\mathbf{0}}$. Any open set is a countable intersection of closed sets, and every closed set is a countable union of open sets, so therefore every open set is in $\sum_{\mathbf{2}}^{\mathbf{0}}$ and every closed set is in $\prod_{\mathbf{2}}^{\mathbf{0}}$. However, for any closed set$B$, $\bigcup_{n=1}^{\infty} B \in \sum_{\mathbf{2}}^{\mathbf{0}}$ and so every closed set is in $\sum_{\mathbf{2}}^{\mathbf{0}}$ and every open set is in $\prod_{\mathbf{2}}^{\mathbf{0}}$.

Now take arbitrary $\alpha < \beta$ and assume all the subset relations above hold. Then take $B_\alpha \in \sum_{\boldsymbol{\alpha}}^{\mathbf{0}}$. Then since $B_\alpha \in \prod_{\boldsymbol{\beta}}^{\mathbf{0}}$, $B_\alpha \in \sum_{\boldsymbol{\beta+1}}^{\mathbf{0}}$, Similarly $B_\alpha \in \sum_{\boldsymbol{\beta}}^{\mathbf{0}} \Rightarrow B_\alpha^C \in \prod_{\boldsymbol{\beta}}^{\mathbf{0}} \Rightarrow B_\alpha^C \in \sum_{\boldsymbol{\beta+1}}^{\mathbf{0}} \Rightarrow B_\alpha \in \prod_{\boldsymbol{\beta+1}}^{\mathbf{0}}$. So $\sum_{\boldsymbol{\alpha}}^{\mathbf{0}} \subset \sum_{\boldsymbol{\beta+1}}^{\mathbf{0}}$ and $\sum_{\boldsymbol{\alpha}}^{\mathbf{0}} \subset \prod_{\boldsymbol{\beta+1}}^{\mathbf{0}}$.

Now take $B_\alpha \in \prod_{\boldsymbol{\alpha}}^{\mathbf{0}}$. Then $B_\alpha \in \prod_{\boldsymbol{\beta}}^{\mathbf{0}} \Rightarrow \sum_{\boldsymbol{\beta+1}}^{\mathbf{0}}$. And $B_\alpha \in \sum_{\boldsymbol{\beta}}^{\mathbf{0}} \Rightarrow B_\alpha^C \in \prod_{\boldsymbol{\beta}}^{\mathbf{0}} \Rightarrow B_\alpha^C \in \sum_{\boldsymbol{\beta+1}}^{\mathbf{0}} \Rightarrow B_\alpha \in \prod_{\boldsymbol{\beta+1}}^{\mathbf{0}}$. So $\prod_{\boldsymbol{\alpha}}^{\mathbf{0}} \subset \sum_{\boldsymbol{\beta+1}}^{\mathbf{0}}$ and $\prod_{\boldsymbol{\alpha}}^{\mathbf{0}} \subset \prod_{\boldsymbol{\beta+1}}^{\mathbf{0}}$.

Thus, by transfinite induction, this property holds for all ordinals $\beta > \alpha$, and α was chosen arbitrarily.

Lemma 4.4 *$\sum_{\boldsymbol{\alpha}}^{\mathbf{0}}$ is closed under countable unions and $\prod_{\boldsymbol{\alpha}}^{\mathbf{0}}$ is closed under countable intersections.*

Take a sequence $B_n \in \sum_{\boldsymbol{\alpha}}^{\mathbf{0}}$. Then $B_n = \bigcup_{i=0}^{\infty} B_{i,n}$, $B_{i,n} \in \sum_{\boldsymbol{a}}^{\mathbf{0}}, a < \alpha$. So then $\bigcup_{n=1}^{\infty} B_n = \bigcup_{n=1}^{\infty} \bigcup_{i=1}^{\infty} B_{i,n}$, and we know $N \times N$ is also countable, therefore $\bigcup_{n=1} B_n \in \sum_{\boldsymbol{\alpha}}^{\mathbf{0}}$.

Take sequence $B_n^C \in \prod_{\boldsymbol{\alpha}}^{\mathbf{0}}$. Then $\bigcap_{n=1}^{\infty} B_n^C = (\bigcup_{n=1}^{\infty} B_n)^C$, and since we know $\bigcup_{n=1}^{\infty} B_n \in \sum_{\boldsymbol{\alpha}}^{\mathbf{0}}$, $\bigcap_{n=1}^{\infty} B_n^C \in \prod_{\boldsymbol{\alpha}}^{\mathbf{0}}$.

We are now ready the prove the following.

Theorem 4.5 *Every Borel Set is in $\sum_{\boldsymbol{\alpha}}^{\mathbf{0}}$ or $\prod_{\boldsymbol{\alpha}}^{\mathbf{0}}$ for some $\alpha < \omega_1$.*

If we show that $S = \bigcup_{\alpha<\omega_1} \sum^{\mathbf{0}}_{\boldsymbol{\alpha}} = \bigcup_{\alpha<\omega_1} \prod^{\mathbf{0}}_{\boldsymbol{\alpha}}$ is closed under countable union and intersection, then we will show that it contains every Borel set, since it clearly contains all open sets.

Take a sequence $(B_n) \in S$. For each B_n take α_n to be $\min_\alpha : (B_n \in \sum^{\mathbf{0}}_{\boldsymbol{\alpha}})$. (This minimum exists since the ordinals are well-ordered). Then using the axiom of choice, we can denumerate each α_n, and then using this enumeration we see $\bigcup \alpha_n$ is countable and therefore $\alpha^* = \sup \alpha_n < \omega_1$, therefore $\bigcup B_n \in \sum^{\mathbf{0}}_{\boldsymbol{\alpha^*+1}}$, and so S is closed under countable union.

We have included this proof to demonstrate that the idea of transfinite induction is not an idea on the fringe. While it is a somewhat fantastical idea (in some ways it is performing an operation 'more than infinity times'), it is something that very much can come up in even elementary analysis.

5 Part 2: Lagarias's Inequality

The aim of this second portion of the project is to provide a thorough exposition of the proofs of Guy Robin and Jeff Lagarias which led to the remarkable proof that the inequality:

$$\sum_{d|n} d \leq e^{H_n} \log(H_n), n \geq 60$$

is equivalent to the Riemann Hypothesis.

To do so, we will first need to follow Robin's footsteps in proving an important quality of "Colossally Abundant Numbers", which are numbers which are exceptionally 'dense' in the sense that the sum of their divisors is large. Essentially, we will demonstrate that to prove inequalities related to the sum of divisors function, we only need to prove them for the colossally abundant numbers. This will serve us well since we can write colossally abundant numbers as a product of products of primes, which are possible to bound if the Riemann Hypothesis is true. Having done so, we will be equipped to prove Robin's inequality:

$$\sum_{d|n} \leq e^{\gamma} n \log\log n, n \geq 5041$$

And from this inequality we will be able to show Lagarias' claim.

First however, we will provide some important background information on the Riemann Hypothesis and how it leads to bounds on certain functions of prime numbers.

6 The Riemann Hypothesis's Importance for Primes

In this section we will attempt to give a (very) brief introduction to the history of the Riemann Hypothesis and why it allows us to place boundaries on some functions involving prime numbers.

Consider the complex function:

$$\zeta(z) = \sum_{n=1}^{\infty} \frac{1}{n^z}$$

This function was originally studied by Euler, but it was Riemann who demonstrated its importance to the study of prime numbers, beginning with the following lemma. (6; 7; 8)

Lemma 6.1

$$\zeta(z) = \prod_p \frac{1}{1-p^{-z}} = \prod_p \frac{p^z}{p^z-1}$$

This is a very interesting result on its own. As an example of how this function already demonstrates a profound relationship to the primes, let $z = 1$, and we get:

$$\zeta(1) = \sum_{n=1}^{\infty} \frac{1}{n} = \infty$$

Which, since no $\frac{p}{p-1}$ is itself infinite, implies there are infinite terms in the product of our second form, meaning there are infinite primes [1].

Importantly, Riemann proved this function has zeros at the negative even integers (the trivial zeros), and that every other zero (the non-trivial zeros) lies in the strip $\{z = \beta + it : \beta < 1\}$ (7).

This bound on β plays a crucial role in the proof of the prime number theorem (14):

Theorem 6.2 (Rosser and Schoenfeld 2.19) *Let $\pi(x)$ be the number of primes less than x, then:*

$$\pi(x) \sim \frac{x}{\log x}$$

The proof makes use of the Tchebychev function $\psi(x) = \sum_{p^m \leq x} \log p$, where p is prime and m is a positive integer. Specifically, it can be shown that (7):

$$\psi(x) \sim x, x \to \infty$$

Implies

$$\pi(x) \sim \frac{x}{\log x}, x \to \infty$$

Furthermore, it can be shown that if we let $\psi_1(x) = \int_1^x \psi(t)dt$, $\psi_1(x) \sim \frac{x^2}{2}$ as $x \to \infty$ implies $\psi(x) \sim x$. The heart of the proof is then showing that $\psi_1(x) \sim \frac{x^2}{2}$ which implies our desired similarity for $\pi(x)$.

Where does the zeta function appear in this? It can be shown that:

$$\psi_1(x) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \frac{x^{t+1}}{t(t+1)} \left(-\frac{\zeta'(t)}{\zeta(t)}\right) dt$$

[1] Not the easiest method to prove this, but an interesting one nonetheless.

This integral can be evaluated where $\zeta(t) \neq 0$, and (to oversimplify), knowing where ζ *cannot* be 0 allows one to 'box around' the zeros and demonstrate that $\psi_1(x) \sim \frac{x^2}{2}$.

Later, an error bound on just *how* large the difference between $\psi_1(x)$ and $\frac{x^2}{2}$ (and therefore between $\pi(x)$ and $\frac{x}{\log x}$) could get was found by de la Vallée Poussin, who used an improved bound on β that said $< 1 - c(\log \gamma)^{-1}$ for some c and γ. This also allowed him to improve the original estimate of π, showing that a better estimate is (9):

$$\pi(x) \sim Li(x) = \int_2^x \frac{1}{\log t} dt = \frac{x}{\log x} + O(\frac{x}{\log^2 x})$$

de la Vallée Poussin's original estimate of the error can be improved if β can be bounded further. For example, it can be shown that for $\frac{1}{2} < \beta < 1$, we have that (10):

$$\pi(x) - Li(x) = O(x^\beta \log x)$$

The famous Riemann hypothesis is that the zeros of the zeta function only exist on the line $\{z = \beta + it : \beta = \frac{1}{2}\}$. This would be the strictest possible bound on . Using this, one can use similar methods to de la Vallée Poussin to further cut down the error estimate. Specifically, if the Riemann Hypothesis is true (ie. $\beta = \frac{1}{2}$) we can say for any epsilon, x can be sufficiently large so that (9):

$$\pi(x) - Li(x) < x^{\frac{1}{2}+\epsilon}$$

This bound on the error of $\pi(x) - Li(x)$ allows us to place bounds on several functions based on the primes. The work specifically cited on this by Robin is from Rosser and Schoenfeld (12; 13), who proved several bounds on the Tchebyshev functions $\theta(x) = \sum_{p \leq x} \log p$ and the $\psi(x)$ from above. Essentially, by limiting the number of primes which can exist below a certain x, it becomes possible to create upper bounds for these sums.

However, before moving directly on to the proof, we must first examine a very specific type of number.

7 Colossally Abundant Numbers

We define a colossally abundant number N as one where there exists an ϵ such that:

$$\forall n \geq 1, \frac{\sigma(n)}{n^{1+\epsilon}} \leq \frac{\sigma N}{N^{1+\epsilon}}$$

Where $\sigma(n) = \sum_{d|n} d$. Essentially, being a colossally abundant number represents, in some sense, having an especially large number of divisors relative to the size of the number.

Erdős and Nicolas proved several important facts about colossally abundant numbers (11):

- If we define:

$$F(x,\alpha) = \log(1 + 1/(x + x^2 + ... + x^\alpha))/\log x$$

for p prime:

$$E_p = \{F(p, \alpha : \alpha \geq 1\}$$
$$E = \bigcup_{p \text{ prime}} E_p$$

- We can denumerate the elements of E such that:

$$E = \{\epsilon_1, \epsilon_2, ..., \epsilon_i, ...\}, i \leq 1 \rightarrow \epsilon_i > \epsilon_{i+1}$$

- For any $\epsilon > 0$, there is a solution of the following equation, which we label x_k:

$$F(x_1, 1) = \epsilon, F(x_k, k) = \epsilon$$

For ease of notation Robin also refers to x_1 as simply x.

Using these definitions, Robin cites the following facts proved by Erdős and Nicolas (11):

- If $\epsilon \notin E$, $\frac{\sigma(n)}{n^{1+\epsilon}}$ attains a unique maximum, N_ϵ which is by definition a colossally abundant number, then prime factorization of this number can be written as:

$$N_\epsilon = \prod_p p^{\alpha_p(\epsilon)}$$

where $\alpha_p(\epsilon)$ is a function defined as:

$$\alpha_p(\epsilon) = \begin{cases} k & x_{k+1} < p < x_k \wedge k \geq 1 \\ 0 & p > x_1 \end{cases}$$

- For every $i \geq 1$, there is a distinct N_i such that for $\epsilon \in (\epsilon_{i-1}, \epsilon)$, $N_\epsilon = N_i$.
- The function $\frac{\sigma(n)}{n^{1+\epsilon_i}}$ obtains a maximum at two distinct points, N_i and N_{i+1}.
- As $\epsilon \rightarrow 0$, $x_k \sim \sqrt{kx_1}$.
- $x_2 = \sqrt{2x_1}(1 - \frac{\log 2}{2\log x_1} + O(\frac{1}{\log^2 x}))$ for $x_1 \geq 1530$

Robin then provides proofs for several lemmas which we simplify and expand.

Lemma 7.1 *For $x_1 > 1, k \geq 2$, $x_k > x^{\frac{1}{k}}$*

We aim to prove that $F(x^{\frac{1}{k}}, k) > \epsilon = F(x, 1)$, letting (for notational cleanliness only) $t = x^{\frac{1}{k}}$ and expanding we get:

$$\frac{\log(1 + 1/(t + t^2 + ... + t^k))}{\log(t)} > \frac{\log(1 + 1/x)}{\log(x)} = \frac{\log(1 + 1/t^k)}{\log(t^k)}$$

Multiplying both sides by $k \log(t)$ we obtain:

$$k \log(1 + 1/(t + t^2 + ... + t^k)) > \log(1 + 1/t^k)$$

Placing the k into the left logarithm and taking exponential gives us:

$$(1 + 1/(t + t^2 + ... + t^k))^k > (1 + 1/t^k)$$

Using the fact that $(1 + x)^k > 1 + kx$ (proved in the appendix), we obtain:

$$(1 + \frac{1}{t + t^2 + ... + t^k})^k > 1 + \frac{k}{t + t^2 + ... + t^k}$$

If $x > 1$ then $t > 1$, and so then $t^j < t^k$ for $j < k$ therefore $t + t^2 + ... + t^k < kt^k$ and so:

$$(1 + \frac{1}{t + t^2 + ... + t^k})^k > 1 + \frac{k}{t + t^2 + ... + t^k} > 1 + \frac{k}{kt^k} = 1 + \frac{1}{t^k}$$

Therefore $F(t, k) = F(x^{\frac{1}{k}}, k) > \epsilon$, and since $F(x, k)$ with k fixed is a decreasing function, $x_k > x^{\frac{1}{k}}$.

Lemma 7.2 *For $x > 1$, $x_2 < \sqrt{2x}$*

We take the same approach as in proof one except instead of showing the value is larger than epsilon, we wish to show that $F(\sqrt{2x}, 2) < \epsilon$, which would imply the lemma. Performing the same expansion as before, we obtain:

$$\frac{\log(1 + \frac{1}{\sqrt{2x}+2x})}{\log(\sqrt{2x})} = \frac{2\log(1 + \frac{1}{\sqrt{2x}+2x})}{\log(x) + \log(2)} < \frac{2\log(1 + \frac{1}{\sqrt{2x}+2x})}{\log(x)}$$

Therefore, multiplying by $\log x$ and performing exponentiation, we see the lemma is proved by showing:

$$(1 + \frac{1}{\sqrt{2x} + 2x})^2 \leq 1 + \frac{1}{x}$$

Multiplying out the left hand side we get:

$$1 + \frac{2}{\sqrt{2x} + 2x} + \frac{1}{4x^2 + 2x + 4x\sqrt{2x}} \leq 1 + \frac{1}{x}$$

Which is equivalent to:

$$\frac{2}{\sqrt{2x} + 2x} + \frac{1}{4x^2 + 2x + 4x\sqrt{2x}} \leq \frac{1}{x}$$

or:

$$\frac{2x}{\sqrt{2x}+2x}+\frac{1}{4x+4\sqrt{2x}+2}=\frac{8x^2+(6+8\sqrt{2x})x+\sqrt{2x}}{8x^2+(12+12\sqrt{2x})x+2\sqrt{2x}}\leq 1$$

Which is clearly true. Therefore, the inequality holds, and so we know that $F(\sqrt{2x},2)<\epsilon$, which implies that $x_2<\sqrt{2x}$.

Lemma 7.3 *For* $x\geq 1530$, $x_2>\sqrt{2x}(1-\frac{\log 2}{2\log x})$

Robin first sets $u=\frac{\log 2}{\log x}$ $z=\sqrt{2x}(1-\frac{u}{2})$, so that our expansion of $F(\sqrt{2x}(1-\frac{\log 2}{2\log x},2)>\epsilon=F(x,1)$ is simplified to:

$$\frac{\log(1+\frac{1}{z+z^2})}{\log z}>\frac{\log(1+\frac{1}{x})}{\log x}$$

Now Robin takes advantage of the following lemma (proved in appendix):

$$\frac{1}{1+a}\leq\log(1+\frac{1}{a})<\frac{1}{a}$$

Applying the above, we see the following inequality is sufficient for the proof:

$$\frac{1}{1+z+z^2}>\frac{\log z}{x\log x}$$

Expanding z we obtain:

$$\frac{\log z}{x\log x}=\frac{1}{x\log x}\log(\sqrt{2x}(1-\frac{u}{2}))=\frac{1}{x\log x}(\frac{1}{2}(\log(2)+\log(x))+\log(1-\frac{u}{2}))$$

$$=\frac{1}{2x\log x}(\log(2)+\log(x)+2\log(1-\frac{u}{2}))=\frac{1}{2x}(\frac{\log 2}{\log x}+\frac{\log x}{\log x}+2\frac{\log(1-\frac{u}{2})}{\log x})$$

$$=\frac{1}{2x}(1+u+2\frac{\log(1-\frac{\log 2}{2\log x})}{\log x})$$

Since $\log(1+b)\leq b, \forall b, -1<b<\infty$ (proved in appendix) and $|\frac{\log 2}{2\log x}|<1$ certainly holds for $x\geq 1530$, we obtain:

$$=\frac{1}{2x}(1+u+2\frac{\log(1-\frac{\log 2}{2\log x})}{\log x})\leq\frac{1}{2x}(1+u-2\frac{\log 2}{2\log x}\frac{1}{\log x})=\frac{1}{2x}(1+u-\frac{\log 2}{\log x^2})$$

$$=\frac{1}{2x}(1+u-\frac{u^2}{\log 2})$$

Therefore we have:

$$\frac{1}{1+z+z^2}>\frac{1}{2x}(1+u-\frac{u^2}{\log 2})$$

$$\Rightarrow (1+z+z^2)(1+u-\frac{u^2}{\log 2}) < 2x$$

Robin proceeds by splitting this sum and evaluating the parts separately. First we evaluate $z^2(1+u-\frac{u^2}{\log 2})$:

$$z^2(1+u-\frac{u^2}{\log 2}) = 2z^2\frac{\log z}{\log x} = 2x\frac{z^2}{x}\frac{\log z}{\log x} = 2x(1-\frac{u}{2})^2(1+u-\frac{u^2}{\log 2})$$

$$= 2x(1-u^2(\frac{3}{4}+\frac{1}{\log 2})+u^3(\frac{1}{4}+\frac{1}{\log 2})-u^4(\frac{1}{4\log 2})) \leq 2x(1-u^2(\frac{3}{4}+\frac{1}{\log 2})+u^3(\frac{1}{4}+\frac{1}{\log 2}))$$

Letting $x > 2^{11}$, $u = \frac{\log 2}{\log x} = \log_x 2 < \frac{1}{11}$, and so $u^3 < \frac{1}{11}u^2$ and we can say:

$$z^2(1+u-\frac{u^2}{\log 2}) < 2x(1-u^2(\frac{3}{4}+\frac{1}{\log 2})+u^2(\frac{1}{11})(\frac{1}{4}+\frac{1}{\log 2})) = 2x(1-u^2(2.193...)+u^2(.154...))$$

$$\leq 2x(1-(2.03)u^2)$$

Now we turn to the $(1+z)(1+u-\frac{u^2}{2})$ part. Multiplying it out we have:

$$(1+z)(1+u-\frac{u^2}{2}) = 1+u-\frac{u^2}{2}+z+uz-z$$

$$\leq (1+u)z+1+u = \sqrt{2x}(1-\frac{u}{2}+u-\frac{u^2}{2})+1+u \leq \sqrt{2x}(1+\frac{u}{2})+1+u$$

Given the same assumptions for x, we know $\sqrt{2x} \geq 64$, and so $1+u \leq 1+\frac{1}{11} = \frac{3}{176}\sqrt{2x}$ and so we obtain:

$$(1+z)(1+u-\frac{u^2}{2}) \leq (1+\frac{u}{2})\sqrt{2}+\frac{3}{176}\sqrt{2} \leq 2.07\sqrt{2x}$$

Putting these parts together we finally obtain:

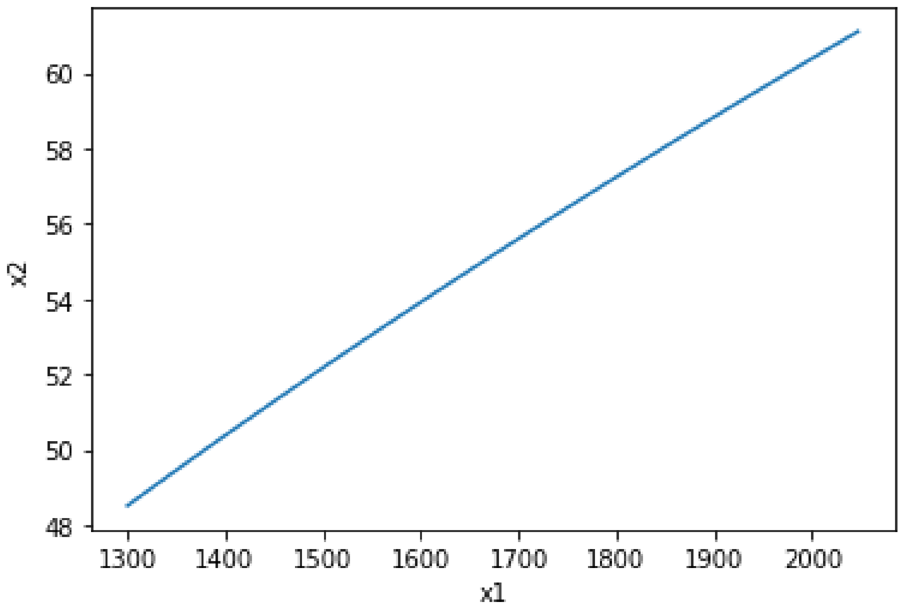
$$\Rightarrow (1+z+z^2)(1+u-\frac{u^2}{\log 2}) \leq 2.07\sqrt{2x}+2x(1-2.03u^2) = 2x+2.07\sqrt{2x}-(4.06u^2)x$$

$$= 2x+2.07\sqrt{2x}-(4.06u^2)x$$

And therefore proving the following inequality is sufficient to complete the proof for the $x \geq 2^{11}$ case:

$$2.07\sqrt{2x} < 4.06x(\frac{\log 2}{\log x})^2$$

Which we know holds based on the growth rate of x versus $\sqrt{x}$. For the $x < 2^{11}$ case, Robin says he checked this by computer. We have replicated his efforts and include the graph below as well as a table of calculated values in the appendix.



8 Robin's First Statement

Let:

$$f(n) = \frac{\sigma(n)}{n \log \log n}$$

Robin's first lemma from his section on his inequality aims to show that for any number n which is between two colossally abundant numbers N and N', $f(n) \leq \max(f(N), f(N'))$.

Lemma 8.1 *For n, N, N' where N and N' are successive colossally abundant numbers, and $3 \leq N \leq n \leq N'$, then $f(n) \leq \max(f(N), f(N'))$.*

From Erdős (11), we know $\exists \epsilon > 0$:

$$\forall n \geq 1, \frac{\sigma(n)}{n^{1+\epsilon}} \leq \frac{\sigma(N)}{N^{1+\epsilon}} = \frac{\sigma(N')}{N'^{1+\epsilon}}$$

Therefore using some algebraic manipulation we get:

$$\frac{\sigma(n)}{n^{1+\epsilon}} \frac{1}{\log \log n} n^{\epsilon} \leq \frac{\sigma(N)}{N^{1+\epsilon}} \frac{1}{\log \log n} n^{\epsilon}$$

$$\Rightarrow \frac{\sigma(n)}{n \log \log n} \leq \frac{\sigma(N)}{N \log \log N} (\frac{n}{N})^{\epsilon} \frac{\log \log N}{\log \log n}$$

$$\Rightarrow f(n) \leq f(N) (\frac{n}{N})^{\epsilon} \frac{\log \log N}{\log \log n}$$

The same proof also gives us that:

$$f(n) \leq f(N')(\frac{n}{N'})^{\epsilon}\frac{\log\log N'}{\log\log n}$$

Therefore we aim to show that:

$$(\frac{n}{N})^{\epsilon}\frac{\log\log N}{\log\log n} \leq 1$$

Which would imply $f(n) \leq f(N)$, or show the equivalent for N'. This inequality is equivalent to:

$$\frac{n^{\epsilon}}{\log\log n} \leq \frac{N^{e}}{\log\log N}$$

Which, taking the natural log of each side, is equivalent to:

$$\epsilon \log n - \log\log\log n \leq \epsilon \log N - \log\log\log N$$

Therefore, we want to show that:

$$\epsilon \log n - \log\log\log n \leq \max(\epsilon \log N - \log\log\log N, \epsilon \log N' - \log\log\log N')$$

As then the above inequality would hold for N or N', indicating that $f(n) \leq f(N) \vee f(n) \leq f(N')$, and thus $f(n) \leq \max(f(N), f(N'))$.

To prove this, let $g(x) = \epsilon x - \log\log x$ for $x > 0$. Then:

$$g'(x) = \epsilon - \frac{x}{\log x}$$

$$g''(x) = \frac{\log x + 1}{(x \log x)^2}$$

And so clearly for $x > 0, g''(x) > 0$, and thus g is convex. Therefore, since log is a positive increasing function, for $N \leq n \leq N'$, $g(\log n) \leq \max(g(\log N), g(\log N'))$. This result is equivalent to the necessary inequality.

This lemma provides the groundwork for the remainder of Robin's proof. It essentially means that we can focus only on proving Robin's inequality for colossally abundant numbers, which can be written in a very particular way.

Theorem 8.2 *If the Riemann Hypothesis is true, there exists an n_0 s.t. $f(n) < e^{\gamma}$ for $n \geq n_0$.*

Robin starts off by using an unusual formula for $\frac{\sigma(N)}{N}$ if N is a colossally abundant number.

As a reminder, in Section 2 Robin cited a result from Erdős and Nicolas (11) that states N can be written as:

$$N = \prod_{x_2 < p \leq x_1} p \prod_{x_3 < p \leq x_2} p^2 \prod_{x_4 < p \leq x_3} p^3 ...$$

Furthermore, number theory states that if n has prime factorization $n = p_1^{k_1} p_2^{k_2} ... p_r^{k_r}$, we can write $\sigma(n)$ as(**?**):

$$\sigma(n) = \prod_{i=1}^{r} (1 + p_i + p_i^2 + ... + p_i^{k_i})$$

And since we know the prime factorization of N, we can write:

$$\sigma(N) = \prod_{x_2 < p \le x_1} (1 + p) \prod_{x_3 < p \le x_2} (1 + p + p^2)...$$

Combining these two results, we see:

$$\frac{\sigma(N)}{N} = \frac{\prod_{x_2<p\le x_1}(1+p)\prod_{x_3<p\le x_2}(1+p+p^2)\prod_{x_4<p\le x_3}(1+p+p^2+p^3)...}{\prod_{x_2<p\le x_1} p \prod_{x_3<p\le x_2} p^2 \prod_{x_4<p\le x_3} p^3...}$$

Then, factoring the corresponding products, we obtain Robin's starting equality:

$$\frac{\sigma(N)}{N} = \prod_{x_2 < p \le x_1} (1 + \frac{1}{p}) \prod_{x_3 < p \le x_2} (1 + \frac{1}{p} + \frac{1}{p^2})...$$

We now use the limit of a geometric series to state that:

$$(1 + \frac{1}{p} + \frac{1}{p^2} + ...) \le \frac{1}{1 - \frac{1}{p}}$$

to compact the terms for $p \le x_2$ and obtain:

$$\frac{\sigma(N)}{N} \le \frac{\le \prod_{x_2<p\le x_1}(1+\frac{1}{p})}{\prod_{p\le x_2}(1-\frac{1}{p})}$$

We also use the fact that:

$$1 - \frac{1}{p^2} = (1 + \frac{1}{p})(1 - \frac{1}{p})$$

To rewrite the numerator as:

$$\prod_{x_2 < p \le x_1} (1 + \frac{1}{p}) = \frac{\prod_{x_2<p\le x_1}(1-\frac{1}{p^2})}{\prod_{x_2<p\le x_1}(1-\frac{1}{p})}$$

From here we get:

$$\frac{\sigma(N)}{N} \le \frac{\prod_{x_2<p\le x_1}(1-\frac{1}{p^2})}{\prod_{p\le x}(1-\frac{1}{p})}$$

Which since we showed $x_2 \le \sqrt{2x_1}$:

$$\frac{\sigma(N)}{N} \le \frac{\prod_{\sqrt{2x}<p\le x_1}(1-\frac{1}{p^2})}{\prod_{p\le x}(1-\frac{1}{p})}$$

Robin then proceed to attempt to bound every part of this formula, as well as bounding $\log\log N$, in order to ultimately bound $f(N)$.

To bound the numerator we can write:

$$\prod_{\sqrt{2x}<p<x}(1-\frac{1}{p^2}) = \prod_{\sqrt{2x}<p<x}\exp(\log(1-\frac{1}{p^2})) = \exp(\sum_{\sqrt{x}<p<x}\log(1-\frac{1}{p^2}))$$

Then using our identity $\log(1-\frac{1}{a}) \leq \frac{1}{a}$ for $-\frac{1}{a} > -1$:

$$\exp(\sum_{\sqrt{x}<p<x}\log(1-\frac{1}{p^2})) \leq \exp(-\sum_{\sqrt{x}<p<x}\frac{1}{p^2})$$

and so we have:

$$\prod_{\sqrt{2x}<p<x}(1-\frac{1}{p^2}) = \exp(-\sum_{\sqrt{x}<p<x}\frac{1}{p^2} + O(\frac{1}{x^{3/2}}))$$

Now Robin uses an uncited proposition about the sum $\sum_{p\geq x}\frac{1}{p^2}$:

Proposition 1

$$\sum_{p\geq x}\frac{1}{p^2} = \frac{1}{x\log x} + O(\frac{1}{x\log^2 x})$$

Using this, we can evaluate (assuming $x > 2$):

$$\sum_{\sqrt{2x}<p<x}\frac{1}{p^2} = \sum_{p\geq\sqrt{2x}}\frac{1}{p^2} - \sum_{p\geq x}\frac{1}{p^2}$$

$$= \frac{1}{\sqrt{2x}\log\sqrt{2x}} + O(\frac{1}{\sqrt{2x}\log^2\sqrt{2x}}) - \frac{1}{x\log x} + O(\frac{1}{x\log^2 x})$$

$$= \frac{\sqrt{2}}{\sqrt{x}(\log 2+\log x)} + O(\frac{2\sqrt{2}}{\sqrt{x}(\log^2 x + 2\log x\log 2 + \log^2 2)}) - \frac{1}{x\log x} + O(\frac{1}{x\log^2 x})$$

For x large enough we have $(\log 2 + \log x) \sim \log x$ and $(\log^2 x + 2\log x\log 2 + \log^2 2) \sim \log^2 x$ (the difference can be absorbed into the O terms), and so rewriting we get:

$$\frac{\sqrt{2}}{\sqrt{x}\log x} - \frac{1}{x\log x} + O(\frac{1}{\sqrt{x}\log^2 x})$$

But since $\frac{1}{x\log x} < \frac{1}{\sqrt{x}\log^2 x}$ for very large x, it too can be absorbed into the error term giving us finally:

$$\sum_{\sqrt{2x}<p<x}\frac{1}{p^2} = \frac{\sqrt{2}}{\sqrt{x}\log x} + O(\frac{1}{\sqrt{x}\log^2 x})$$

And so therefore:

$$\prod_{\sqrt{2x}<p<x}(1-\frac{1}{p^2}) = \exp(-\sum_{\sqrt{x}<p<x}\frac{1}{p^2}+O(\frac{1}{x^{3/2}})) = \exp(-\frac{\sqrt{2}}{\sqrt{x}\log x}+O(\frac{1}{\sqrt{x}\log^2 x}))$$

In order to bound $\log\log N$, we first remember that from Erdős and Nicolas (11) we can write:

$$N = \prod_{x_2<p\leq x_1} p \prod_{x_3<p\leq x_2} p^2...$$

Re-arranging, we see this is the same as:

$$N = \prod_{p\leq x_1} p \prod_{p\leq x_2} p \prod_{p\leq x_3} p...$$

And so by taking the log we can split this into:

$$\log N = \sum_{p\leq x_1}\log p + \sum_{p\leq x_2}\log p + ...$$

$\sum_{p\leq x}\log p$ is a well-studied function called the Chebyshev prime function, and is usually denoted $\theta(x)$. Since $\sum_{p\leq x_n}\log p \leq \sum_{p\leq x_3}\log p$ for all $n\geq 3$, we can say:

$$\log N = \theta(x_1) + \theta(x_2) + O(x_3)$$

$$\log N = \theta(x_1)(1+\frac{\theta(x_2)}{\theta(x_1)}+\frac{O(x_3)}{\theta(x_1)}) = \theta(x_1)(1+\frac{\sqrt{2}}{\sqrt{x}}+O(\frac{1}{\sqrt{x}\log x}))$$

$$\log\log N = \log\theta(x)\exp(\frac{\sqrt{2}}{\sqrt{x}\log x}+O(\frac{1}{x\log^2 x}))$$

Finally, Robin cites a result from Nicolas (15), which states that:

$$\prod_{p\leq x}(1-\frac{1}{p}) \geq \frac{e^{-\gamma}}{\log\theta(x)}\exp(\frac{-(2+c)}{\sqrt{x}\log x}+O(\frac{1}{\sqrt{x}\log^2 x})$$

where γ is the Euler-Mascheroni constant which is equal to .57721... and c is the result of the Prime Zeta function, $\sum_p \frac{1}{|p|^2}$. Robin cites a result from H.M Edwards (9) which tells us that:

$$c = \gamma + 2 - \log 4\pi$$

So finally, substituting all of these results, we obtain:

$$f(n) = \frac{\sigma(n)}{\log\log N} \leq \frac{\prod_{\sqrt{2x}<p\leq x}(1-\frac{1}{p^2})}{\prod_{p\leq x}(1-\frac{1}{p})(\log\log N)}$$

$$\leq \exp(\frac{-\sqrt{2}}{\sqrt{x}\log x}+O(\frac{1}{\sqrt{x}\log^2 x}))\frac{1}{\log\theta(x)\exp(\frac{\sqrt{2}}{\sqrt{x}\log x}+O(\frac{1}{x\log^2 x}))}\frac{\log\theta(x)}{e^{-\gamma}}\frac{1}{\exp(\frac{-(2+c)}{\sqrt{x}\log x}+O(\frac{1}{\sqrt{x}\log^2 x}))}$$

Combining the exponents and performing cancellations we get:

$$= e^{\gamma} \exp(\frac{2 + c - 2\sqrt{2}}{\sqrt{x}\log x} + O(\frac{1}{\sqrt{x}\log^2 x}))$$

Evaluating $2 + c - 2\sqrt{2}$ as $-.782...$, it is clear that for x sufficiently large, (therefore N sufficiently large):

$$\frac{2 + c - 2\sqrt{2}}{\sqrt{x}\log x} + O(\frac{1}{\sqrt{x}\log^2 x}) < 0$$

Therefore for N suffieciently large:

$$\exp(\frac{2 + c - 2\sqrt{2}}{\sqrt{x}\log x} + O(\frac{1}{\sqrt{x}\log^2 x})) < 1$$

and finally this implies:

$$f(N) < e^{\gamma}$$

9 Finding n_0

Having demonstrated that there is an n_0 for which this statement holds, Robin goes about proving that this n_0 is in fact 5041. To do so, he uses several bounds (dependent on the Riemann hypothesis) which he derives from Rosser and Schoenfeld's bounds on the Chebyshev functions (12; 13) combined computer calculation.

- For $x \geq 20000$:

$$\prod_{\sqrt{2x} < p \leq x} (1 - \frac{1}{p^2}) \leq \exp(\frac{-\sqrt{2}}{\sqrt{x}\log x} + \frac{4}{\sqrt{x}\log^2 x})$$

- For $x \geq 20000$:

$$\log\log N > \log\theta(x)\exp(\frac{0.968\sqrt{2}}{\sqrt{x}\log x} - \frac{0.342}{\sqrt{x}\log^2 x})$$

- For $x \geq 20000$:

$$\prod_{p \leq x}(1 - \frac{1}{p})^{-1} \leq e^{\gamma}\log\theta(x)\exp(\frac{2 + c}{\sqrt{x}\log x} + \frac{\alpha(x)}{\sqrt{x}\log^2 x})$$

  Where α is defined as:

$$\alpha = \frac{(\theta(x) - x)^2(\log x + 1.31)}{2x^{\frac{3}{2}}} + (c - 2) + \frac{8 + 4c}{\log x} + \frac{2\log x}{x^{\frac{1}{6}}} + \frac{\log 2\pi \log x}{x^{\frac{1}{2}}}$$

These bounds are easily recognized as sharper bounds on the components of the upper bound of $\frac{\sigma(N)}{N}$, which we used in the proof above. With them, the following statement becomes possible:

Theorem 9.1 *If the Riemann Hypothesis is true, $\sigma(n) < e^\gamma n \log\log n$ for $\forall n \geq$* 5041

We showed in section 3 that for a colossally abundant number N:

$$f(N) = \frac{\sigma(N)}{N \log\log N} \leq \prod_{x_2 < p \leq x} (1 - \frac{1}{p^2}) \prod_{p \leq x} (1 - \frac{1}{p})^{-1} (\log\log N)^{-1}$$

Substituting the bounds above in, we obtain:

$$f(N) \leq \frac{(\exp(\frac{-\sqrt{2}}{\sqrt{x}\log x} + \frac{4}{\sqrt{x}\log^2 x}))(e^\gamma \log\theta(x) \exp(\frac{2+c}{\sqrt{x}\log x} + \frac{\alpha(x)}{\sqrt{x}\log^2 x}))}{(\log\theta(x) \exp(\frac{0.968\sqrt{2}}{\sqrt{x}\log x} - \frac{0.342}{\sqrt{x}\log^2 x}))}$$

$$= e^\gamma \exp(\frac{2 + c - 1.968\sqrt{2}}{\sqrt{x}\log x} + \frac{\alpha(x) + 4.342}{\sqrt{x}\log^2 x})$$

$$= e^\gamma \exp(\frac{1}{\sqrt{x}\log x}(2 + c - 1.968\sqrt{2} + \frac{\alpha(x) + 4.342}{\log x}))$$

Since we want the exponent of the second e to be negative, we must show that:

$$\frac{\alpha(x) + 4.342}{\log x} < 2 + c - 1.968\sqrt{2} = .7624...$$

Robin now breaks the proof into three parts, showing this relation is true for $x \geq 100000$, $20000 \leq x \leq 100000$, and $x \leq 20000$. First, in proving the $x \geq 100000$ case, he cites a result from Schoenfeld's "Sharper Bounds" paper (12):

$$\forall x \geq 599, |\theta(x) - x| \leq \frac{1}{8\pi}\sqrt{x}\log^2 x$$

Plugging this into $\alpha(x)$ we get:

$$\alpha(x) \leq \frac{\log^4 x(\log x + 1.31)}{128\pi^2 x^{\frac{1}{2}}} + (c - 2) + \frac{8 + 4c}{\log x} + \frac{2\log x}{x^{\frac{1}{6}}} + \frac{\log 2x \log x}{x^{\frac{1}{2}}}$$

While this equation looks complicated, we can see that the denominators all feature powers of x while the numerators are composed of logs, and thus eventually the former will dominate the latter. Calculating this sum at $x = 100000$, we get that $\alpha(100000) \leq 3.145...$, so $\alpha(x) < 3.15, x \geq 100000$. Plugging this into our previous equation gives us:

$$\forall x \geq 100000, \frac{\alpha(x) + 4.342}{\log x} \leq \frac{3.15 + 4.342}{\log 100000} = .6507... < .7624$$

Thus proving our inequality for $x \geq 100000$.

Next Robin takes two inequalities from the same Schoenfeld paper (**?**):

- $\forall x \leq 10^{11}, \theta(x) < x$ [p. 360]
- $\forall x \geq 19421, \theta(x) > x - \frac{x}{8 \log x}$ [p. 359]

Combining these, we see that for $20000 \leq x \leq 100000$, we can bound $|\theta(x) - x| \leq |\frac{x}{8 \log x}|$, and putting this into our equation for α we get:

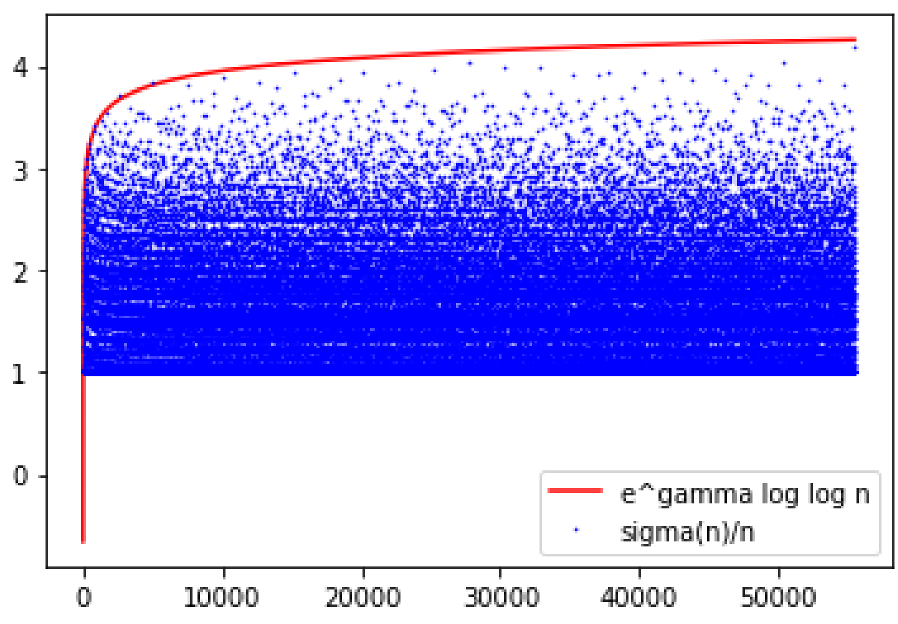
$$\alpha(x) \leq \frac{\sqrt{x}(\log x + 1.31)}{128 \log^2 x} + (c-2) + \frac{8+4c}{\log x} + \frac{2 \log x}{x^{\frac{1}{6}}} + \frac{\log 2\pi \log x}{x^{\frac{1}{2}}}$$

This is an easily calculated function, and so calculating the values for $20000 \leq x \leq 100000$ on the computer, we see that the function never goes above the value it obtains at $\alpha(20000) = 2.92938....$ Therefore $20000 \leq x \leq 100000 \rightarrow \alpha(x) < 3$ and so:

$$\frac{\alpha(x) + 4.343}{\log x} \leq \frac{3 + 4.342}{\log 20000} = .7414... < .7624$$

Finally, for $x \leq 20000$, Robin simply calculates each Colossally Abundant number and confirms it meets the desired inequality. We have replicated his efforts, which show that for each such number above 5040, starting with 55440, the inequality holds. Therefore, by Lemma 7.1 we have that this holds for $n \geq 55440$.

To prove for $n \leq 55440$ we once again turn to the computer (as Robin also did), and we find that this inequality holds for any $n \geq 5041$ as expected. The chart of $\frac{\sigma(n)}{n}$ vs $e^{\gamma} \log \log n$ is provided below. The full results for both of these computations are provided in Appendix 2 and 3 respectively.



Having demonstrated these bounds on $\frac{\sigma(n)}{n}$, we are prepared to follow Lagarias's proof that his inequality holds.

10 Lagarias's Inequality

Lagarias uses two facts from Robin (17):

- If the Riemann hypothesis is true, $\forall n \geq 5041$:

$$\sigma(n) < e^{\gamma} n \log\log n$$

- If the Riemann hypothesis is false, $\exists 0 < \beta < \frac{1}{2}$ and $C > 0$ such that:

$$\sigma(n) \geq e^{\gamma} n \log\log n + \frac{Cn \log\log n}{(\log n)^{\beta}}$$

  holds for infinitely many n.

He then proves two lemmas of his own, which constitute the meat of the proof.

Lemma 10.1 *For $n \geq 3$,*

$$\exp(H_n)\log(H_n) \geq e^{\gamma} n \log\log n$$

First he begins with the integral:

$$\int_1^n \frac{\lfloor t \rfloor}{t^2} dt$$

Where $\lfloor t \rfloor$ denotes the floor function. We can split this integral up into:

$$\int_1^2 \frac{\lfloor t \rfloor}{t^2} dt + \int_2^3 \frac{\lfloor t \rfloor}{t^2} dt + ... + \int_{n-1}^n \frac{\lfloor t \rfloor}{t^2} dt = \sum_{k=1}^{n-1} \int_k^{k+1} \frac{\lfloor t \rfloor}{t^2} dt = \sum_{k=1}^{n-1} k \int_k^{k+1} \frac{1}{t^2} dt$$

Since $\int_a^b \frac{1}{t^2} dt = \frac{1}{a} - \frac{1}{b}$ we have:

$$\sum_{k=1}^{n-1} k(\frac{1}{k} - \frac{1}{k+1}) = \sum_{k=1}^{n-1}(1 - \frac{k}{k+1}) = \sum_{k=1}^{n-1} \frac{1}{k+1} = \sum_{k=1}^{n} \frac{1}{k} - 1$$

And so we have:

$$\int_1^n \frac{\lfloor t \rfloor}{t^2} dt = H_n - 1$$

Lagarias then defines $\{t\} = t - \lfloor t \rfloor$ and rewrites this as:

$$H_n = 1 + \int_1^n \frac{t - \{t\}}{t^2} = 1 + \int_1^n \frac{1}{t} dt - \int_1^n \frac{\{t\}}{t^2} dt = 1 + \log n - \int_1^n \frac{\{t\}}{t^2} dt \quad (1)$$

He defines a constant:

$$\gamma = 1 - \int_1^\infty \frac{\{t\}}{t^2} dt$$

Substituting this into our equation we get:

$$\gamma = H_n - \log n - \int_n^\infty \frac{\{t\}}{t^2} dt$$

Letting $n \to \infty$ we get that:

$$\gamma = \lim_{n\to\infty} (H_n - \log n)$$

And so we see this constant is actually the Euler-Mascheroni constant, and so:

$$H_n = \log n + \gamma + \int_n^\infty \frac{\{t\}}{t^2} dt$$

And since $\int_n^\infty \frac{\{t\}}{t^2} dt$ is always non-zero and positive for $n \in N$:

$$H_n > \log n + \gamma$$

$$\Rightarrow \exp(H_n) > \exp(\log n + \gamma) = e^\gamma n$$

Then Lagarias uses the fact that:

$$H_n = \sum_{r=1}^{n} \frac{1}{r}$$

is in fact an upper Riemann sum for $\int_1^n \frac{1}{t} dt = \log n$, and therefore:

$$H_n \geq \log n \Rightarrow \log(H_n) \geq \log \log n$$

Combined with the other result, we obtain:

$$\exp(H_n) \log(H_n) \geq e^\gamma n \log \log n$$

Lemma 10.2 *For $n \geq 3$, $H_n + \exp(H_n) \log(H_n) \leq e^\gamma n \log \log n + \frac{4n}{\log n}$.*

Lagarias proves this starting with the quantity:

$$R_n = H_n - \log(n+1)$$

He observes that both of these can be rewritten as:

$$R_n = H_n - \log(n+1) = \int_1^{n+1} \frac{1}{\lfloor t \rfloor} dt - \int_1^{n+1} \frac{1}{t} dt$$

Since $\frac{1}{[t]} \geq \frac{1}{t}$, we get that this sequence is positive and monotonically increasing, therefore since:

$$\lim_{n\to\infty}(H_n - \log(n+1)) = \gamma$$

by definition, we see that:

$$H_n - \log(n+1) \leq \gamma$$
$$\Rightarrow \exp(H_n)(n+1)^{-1} \leq e^{\gamma}$$
$$\exp(H_n) \leq e^{\gamma}(n+1)$$

From Eq. 1 in the proof of lemma 1 we demonstrated that :

$$H_n = 1 + \log n - \int_1^n \frac{\{t\}}{t^2} dt$$

$$\Rightarrow H_n < \log n + 1 = \log n(1 + \frac{1}{\log n})$$

$$\Rightarrow \log H_n < \log\log n + 1 = \log(\log n(1 + \frac{1}{\log n}) = \log(\log n) + \log(1 + \frac{1}{\log n})$$

Then by the log inequality shown in the appendix:

$$\log(\log n) + \log(1 + \frac{1}{\log n}) \leq \log\log n + \frac{1}{\log n}$$

$$\Rightarrow \log(H_n) \leq \log\log n + \frac{1}{\log n}$$

Multplying these together, Lagarias obtains the formula:

$$\exp(H_n)\log(H_n) \leq e^{\gamma} n \log\log n + \frac{e^{\gamma} n}{\log n} + e^{\gamma}(\log\log n + \frac{1}{\log n})$$

Now Lagarias uses the fact that for $n \geq 3$, $\log\log n + \frac{1}{\log n} \leq \frac{n}{2\log n}$, substituted into the above we get:

$$\exp(H_n)\log(H_n) \leq e^{\gamma} n \log\log n + \frac{e^{\gamma} n}{\log n} + \frac{e^{\gamma} n}{2\log n} = e^{\gamma} n \log\log n + \frac{3e^{\gamma} n}{2\log n}$$

Then using our fact that $H_n \leq \log(n+1) \leq \frac{n}{\log n}$ we get:

$$H_n + \exp(H_n)\log(H_n) \leq e^{\gamma}\log\log n + \frac{(1 + \frac{3}{2}e^{\gamma})n}{\log n} < e^{\gamma}\log\log n + \frac{4n}{\log n}$$

Theorem 10.3 *The statement that* $\forall n$, $\sigma(n) \leq H_n + \exp(H_n)\log(H_n)$*is equivalent to the Riemann Hypothesis.*

If the Riemann Hypothesis is true, then the first cited result from Robin gives us that:

$$\sigma(n) < e^{\gamma} n \log\log n \leq H_n + \exp(H_n)\log(H_n)$$

Which means the Riemann Hypothesis implies our statement.

If our statement is true, then we have that:

$$\sigma(n) \leq H_n + \exp(H_n)\log(H_n) \leq e^{\gamma}\log\log n + \frac{4n}{\log n}$$

Since $\log\log n \to \infty$, we have that for every $C > 0$, $\exists N_C$ such that for $N > N_C$:

$$C \log\log N > 4$$

We also have that for $n \geq 3$, $0 < \beta < \frac{1}{2}$:

$$(\log n)^{\beta} \leq \log n$$

$$\Rightarrow \frac{1}{(\log n)^{\beta}} \geq \frac{1}{\log n}$$

Therefore, combining these, we see that for all $C > 0$, $0 < \beta < \frac{1}{2}$, there exists N_C such that for $n \geq N_C$:

$$\frac{Cn\log\log n}{(\log n)^{\beta}} > \frac{4n}{\log n}$$

$$\Rightarrow \sigma(n) \leq e^{\gamma} n \log\log n + \frac{Cn\log\log n}{(\log n)^{\beta}}$$

Since our second cited proposition from Robin states that if the Riemann Hypothesis was false, there would be infinitely many n which violate the above inequality, our statement must imply the Riemann Hypothesis.

11 Appendix: Miscellanious Lemmas

Robin uses several (easily demonstrated) lemmas without proof. When including the proof within the context the lemma is used would overly slow down the exposition, the proof is included here instead.

Lemma 11.1 *For $k \geq 1, \forall x > 0$, $(1+x)^k \leq 1 + kx$.*

This is trivially true for $k = 1$: $1 + x \geq 1 + x$. Now take the inductive hypothesis that $(1+x)^k \geq 1 + kx$. Then:

$$(1+x)^k \geq 1 + kx$$

$$\Rightarrow (1+x)^{k+1} \geq (1+kx)(1+x) = 1 + kx^2 + kx + x = 1 + kx^2 + (k+1)x$$

$$kx^2 > 0 \Rightarrow (1+x)^{k+1} \geq 1 + kx^2 + (k+1)x \geq 1 + (k+1)x$$

And so we have proved the theorem by induction.

Lemma 11.2

$$-1 < x < \infty \rightarrow \log(1+x) < x$$

Consider $g(x) = \log(1+x) - x$, then we have:

$$g(0) = 0$$

and:

$$g'(x) = \frac{1}{1+x} - 1 = \frac{1}{1+x} - \frac{1+x}{1+x} = \frac{-x}{1+x}$$

and so clearly for $x \in (-1, 0)$, $g'(x) > 0$, and for $x \in (0, \infty), g'(x) < 0$. Combining these results we get:

$$-1 < x < \infty, g(x) \leq g(0) = 0$$

$$\Rightarrow \log(1+x) \leq x$$

Lemma 11.3

$$\frac{1}{1+a} \leq \log(1+\frac{1}{a}) \leq \frac{1}{a}$$

The right inequality comes directly from Lemma 4.2 so we only prove the left. First define $g(x) = \log(1+x) - \frac{x}{1+x}$, and so:

$$g(0) = 0$$

$$g'(x) = \frac{1}{1+x} - \frac{1+x-x}{(1+x)^2} = \frac{1}{1+x} - \frac{1}{(1+x)^2}$$

And so for $x > 0$, we see $g'(x) > 0$, so $g(x) > 0$ and:

$$\log(1+x) \geq \frac{x}{1+x}$$

Letting $x = \frac{1}{a}$, we see $x > 0 \iff a > 0$ and substituting in we obtain:

$$\frac{\frac{1}{a}}{1+\frac{1}{a}} = \frac{1}{a(1+\frac{1}{a})} = \frac{1}{1+a} < \log(1+\frac{1}{a})$$

References

[1] Nicolas Bourbaki. *The Theory of Sets*, Springer. 2004

[2] Halmos, Paul R. *Naive Set Theory*. Springer, 1987.

[3] *Σ-algebra*. In Wikipedia. https://en.wikipedia.org/wiki/%CE%A3-algebra

[4] *Borel Hierarchy*. In Wikipedia. https://en.wikipedia.org/wiki/Borel_hierarchy

[5] Hrbacek, Karel, and Thomas J. Jech. *Introduction to Set Theory*. Marcel Dekker, 1999.

[6] *The Riemann Hypothesis*. In PrimePages, University of Tennessee at Martin. https://primes.utm.edu/notes/rh.html

[7] Riffer-Reinhart, Ben. *The Zeta Function and its Relation to the Prime Number Theorem*. University of Chicago. https://www.math.uchicago.edu/ may/VIGRE/VIGRE2011/REUPapers/Riffer-Reinert.pdf

[8] *Riemann Zeta Function*. In Wikipedia. https://en.wikipedia.org/wiki/Riemann_zeta_function#Definition

[9] Edwards, H. M. *Riemann's Zeta Function*. New York, 1974.

[10] Ingham, A.E. *The Distribution of Prime Numbers*, Cambridge Tracts in Mathematics and Mathematical Physics, Cambridge University Press, 1932. Reprinted 1990.

[11] Erdős, Paul and Nicolas, Jean-Louis. *Répartition des nombres superabondants*. Bulletin de la S.M.F., tome 103, p. 65-90, 1975.

[12] Rosser, J. Barkley, and Schoenfeld, Lowell. Sharper Bounds for the Chebyshev Functions $\theta(x)$ and $\psi(x)$. I. Mathematics of Computation, Volume 29, Number 129, p. 243-269, 1975.

[13] Schoenfeld, Lowell. *Sharper Bounds for the Chebyshev Functions $\theta(x)$ and $\psi(x)$. II.* Mathematics of Computation, Volume 30, Number 134, p. 337-360, 1976.

[14] Rosser, J. Barkley, and Schoenfeld, Lowell. Approximate Formulas for Some Functions of Prime Numbers. Illinois Journal of Math, p. 64-94, 1962.

[15] Nicolas, J. L. *Petites valeurs de la fonction d'Euler*. Journal of Number Theory, Volume 17, Number 3, p. 375-388, 1983.

[16] Robin, Guy. *Grandes Valeurs De La Fonction Somme Des Diviseurs Et Hypotheses De Riemann*. Journal De Mathematiques Pures et Appliquees, Tome 63, Fascicule 2, p.187-213, 1984.

[17] Lagarias, Jeffrey. An Elementary Problem Equivalent to the Riemann Hypothesis, July 29, 2001.

[18] Noe, T.D. *Table of n, a(n) for n = 1...10000* In OEIS. https://oeis.org/A073751/b073751.txt

[19] Oudard, Christian. *Print extremely large long in scientific notation* Stack Overflow. February 25, 2010. https://stackoverflow.com/a/2330992